Incorporating CMIP5 Precipitation Projections into IDF Estimates for the Kansas City Area

By

Madison Crowl

B.S., University of Arkansas, 2017

Submitted to the graduate degree program in Civil, Environmental and Architectural Engineering and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Master of Science in Environmental Engineering.

_____

Chair: Dr. Joshua Roundy

_____

Dr. Edward Peltier

_____

Dr. Bryan Young

_____

Date Defended: 15 November, 2019

The thesis committee for Madison Crowl certifies that this is the approved version of the following thesis:

Incorporating CMIP5 precipitation projections into IDF estimates for the Kansas City area

Chair: Dr. Joshua Roundy

Date Approved: 19 December 2019

Abstract

There is building evidence that climate change is causing an intensification of precipitation patterns. Locations around the world can expect to experience more intense precipitation events. Engineers must be able to account for this future climatic uncertainty in their designs in order to develop sustainable and resilient systems. Intensity-Duration-Frequency (IDF) estimates, developed by the National Oceanic and Atmospheric Administration (NOAA), are used across the United States for engineering design. These estimates were developed under the assumption of a stationary climate (with respect to precipitation intensity). However, research has shown that this assumption may lead to the underestimation of extreme events. The objective of this study is to characterize projected changes in storm intensity, duration, and frequency for the Kansas City area, including: 1) Identifying precipitation trends and 2) Developing IDF estimates incorporating projected climate trends. To achieve these objectives, precipitation data was analyzed from six NOAA gages and the Coupled Model Intercomparison Project version 5 (CMIP5) ensemble. Annual and monthly precipitation was analyzed for both the observational gage data and CMIP5 model data using the Mann Kendall trend test. Increasing trends were identified in the winter (Dec.-Feb.) and spring (Mar.-May) months, while decreasing trends were identified for July-September, indicating a potential shift in seasonal precipitation patterns. Increasing trends were identified for annual precipitation for both the gage data and climate model data. Partial Duration Series (PDS) were developed for the six gages using the Peak-over-threshold (POT) method. Significant increasing trends were identified for the frequency of PDS events. A strong correlation was identified between PDS event frequency and annual precipitation. This relationship was used to develop a novel approach for incorporating climate model projections at the monthly scale into gage-based PDS events used to derive IDF curves. In this methodology, the PDS annual exceedance rate for the future time period was determined based on the CMIP5 model projected annual precipitation. IDF estimates incorporating projected climate model trends were then developed using the adjusted PDS data sets. Results showed an increase in event magnitude from the original for most durations and recurrence

intervals, across all gages, with the 2-year and 100-year events increasing the most. The increase in the event magnitude has serious implications for engineering design. Critical infrastructure, such as bridges, roads, and overflow channels, that are designed for using stationary methods may be under designed.

Acknowledgements

I would like to express my appreciation and thanks to my advisor, Dr. Joshua Roundy, who's guidance and support was invaluable throughout the work. I would like to thank my committee members, Dr. Edward Peltier and Dr. Bryan Young. I would also like to thank my friends and family, and especially my parents, for their love and support throughout.

Table of Contents

List of Tables

List of Figures

List of Equations

Chapter 1: Introduction

A changing climate has significant impacts on the global water cycle. Along with increasing temperatures comes an increase in the variability of precipitation. In 2007, the Intergovernmental Panel on Climate Change (IPCC) determined that climate change effects were already prevalent in the frequency, intensity, and duration of extreme precipitation and droughts (Mondal & Mujumdar, 2015). Both climate models and observed trends have shown an intensification of extreme events (Dore, 2005; Pendergrass & Knutti, 2018). With our current trajectory, this increase in the magnitude and frequency of extreme events is only expected to accelerate, along with altered natural variability (Sarhadi & Soulis, 2017).

Engineers rely on accurate precipitation estimates. The design of water-related infrastructure is only as good as the climate information it employs. Engineers must be able to account for future climatic uncertainty in their designs in order develop sustainable systems. In order to do so, many important aspects must be considered.

An understanding of the changing hydrologic cycle and precipitation patterns is key to designing more resilient water resource systems. Historical information, as well as modeled climate data, can be used to develop a better knowledge of these patterns at both a global and local scale. Hydrologic statistical methods must also be able to capture these patterns to produce improved outputs for engineering design. Within the framework of a changing climate, it is imperative that designs consider all these factors, and not only historical data. Each of these factors is discussed below.

1.1 Climate Observations

Increasing surface temperatures are expected to accelerate the hydrological cycle. The likely changes in both the atmospheric circulation and increases in the atmospheric water-holding capacity will

cause changes in the nature of precipitation (Dore, 2005). While global mean precipitation is expected to increase modestly (~2% K^{-1}), expected increases in extreme precipitation far exceed that (>6% K^{-1}) (Pendergrass & Knutti, 2018). Globally, the rate of increase for precipitation is greater than the evapotranspiration rate, because increased levels of CO_2 in the atmosphere have adverse effects on plant transpiration (Oki & Kanae, 2006).

Evidence shows that moist, energy-limited regions, are becoming wetter, and dry, moisture-limited regions are becoming drier (Allan & Soden, 2008; Prein et al., 2017). Although the hydrological cycle is accelerating, the changing seasonal patterns and the intensification of extreme events can lead to both wetter and drier conditions in the same region (Oki & Kanae, 2006). The altered frequency and intensity of precipitation can mean larger quantities of rain, but also longer dry periods.

Seasonal and inter-annual variabilities, as well as geographic distributions, are also expected to change significantly. Evapotranspiration, precipitation occurring as rain or snow, and snow and ice accumulation and melt could all be strongly influenced by the elevated temperature (Dore, 2005). Higher temperatures will lead to an earlier snowmelt season, causing a shift in spring flood timing and volume (Oki & Kanae, 2006). Climate induced changes in snow directly impacts the Great Plains, as one study detected a 70% decrease of snow cover in the Great Plains (Dore, 2005).

The result of these changing seasonal variabilities is a redistribution of water that impacts soil moisture and groundwater conditions. This will affect the temporal distribution and magnitude of runoff, altering the intensity of both floods and droughts, with streamflows generally increasing in winter and declining in summer (Dore, 2005). This will also affect the quality (due to soil erosion) and distribution of available water. These large shifts in the hydrological cycle will influence the water supply system, sediment transport and deposition, ecosystem conservation, and water-related infrastructure (Jiang et al., 2007).

Engineers and planners manage these systems to reduce flood risk as much as possible. Even so, in the last century in the United States, flooding has caused dozens of fatalities and economic losses in the billions of dollars (Mallakpour & Villarini, 2015). Floods can cause severe and dangerous damage to important infrastructure, such as dams, sewers, and bridges. They also present other hazards, such as flash floods, landslides, and debris flows. Major flood events can threaten food security and displace entire communities. An increase in precipitation, especially extreme precipitation, increases the likelihood of all these risks.

Substantial increases in extreme precipitation in North America have been identified by many studies (Sarhadi & Soulis, 2017). Over the last century, global land precipitation has increased by a statistically significant 2% (Dore, 2005). Regional increases are more variable, with the central United States seeing an even greater increase. When averaged, the global zone from 30N to 85N (the zone including Kansas City), had precipitation increases of 7%-12% (Dore, 2005). Observational data supports a widespread increase in both precipitation intensity and frequency, along with longer dry periods and a higher percentage of precipitation that becomes runoff (Wang et al., 2017). Mallakpour & Villarini (2015) found significant trends in rainfall extremes, with stronger increases in the frequency of extreme events, especially at a seasonal scale. In the Kansas City area, results indicated an increase in both magnitude and frequency annually. Seasonal analysis indicated that increasing magnitude occurred only in the spring and winter, and increasing frequency occurred in the spring, fall, and winter.

Precipitation is, by nature, uneven, and is projected to become even more so. Half of annual precipitation falls in only 12 days of the year (Pendergrass & Knutti, 2018). Pendergrass & Knutt, (2018) defined unevenness by the fraction of total precipitation as a function of the cumulative frequency distribution (33). They found that the percentage of total precipitation that falls during extreme events (defined as the 95^{th} percentile) increased over the last century worldwide, and corresponding fewer days had moderate precipitation. The frequency of extreme precipitation events is increasing more than total precipitation, leading to the decreasing frequency of moderate precipitation events.

Despite substantial evidence of extreme precipitation increasing, corresponding evidence of increases in flooding is lacking (Yue, Pilon, & Cavadias, 2002). Few studies have found evidence of increasing flood magnitude, and some studies have found more evidence of the opposite (Sharma, Wasko, & Lettenmaier, 2018). Globally, the evidence for trends in historical records is inconclusive. The IPCC stated that "there continues to be a lack of evidence and thus low confidence regarding the sign of trend in the magnitude and/or frequency of floods at a global scale" (IPCC, 2013). Multiple studies looking at annual maximum peak discharge magnitude agreed, failing to find any widespread evidence of an increasing trend (Mallakpour & Villarini, 2015). While there is a lack of evidence for global streamflow trends, significant positive streamflow trends have been detected in North America (Šraj et al., 2016). A marked increasing trend in flood damage has been identified in the US (Dore, 2005).

One study by Mallakpour & Villarini (2015) analyzed observational records of 774 stream gages across the US. They identified no statistically significant trends (based on the Mann Kendall Trend test) in the majority of the study area, with 20% that had statistically significant changes, and just 13% with statistically significant increasing trends. Flood frequency, however, showed strong evidence of increasing trends. Of the 774 gages, 34% showed increasing flood frequency, and only 9% showed decreasing flood frequency. Their results showed strong regional differences, most likely due to the integration of climate, stream dynamics, and varying watershed characteristics. For the Kansas City area, they found a decreasing trend in flood frequency overall. However, seasonality was found to be an important aspect. For the summer, an increasing trend in flood frequency was found in the Kansas City area.

These flood trends do not exactly follow the corresponding trends of precipitation. In fact, high flows or flooding are not necessarily driven by individual extreme precipitation events (Šraj, Viglione, Parajka, & Blöschl, 2016). Streamflow depends on more than climatic conditions; other variables, like antecedent soil moisture, have large impacts on streamflow (Šraj et al., 2016). Ivancic & Shaw (2015) found that only 36% of extreme precipitation events corresponded to extreme discharge. For events with

antecedent wet conditions, this number jumps to 62% corresponding to high flows. With antecedent dry conditions, only 13% corresponded to high flows, showing the strong influence of antecedent moisture conditions on streamflow magnitude. However, this influence changes for highly urbanized catchments. The high percentage of impervious surface gives urban catchments low storage, and it can be expected that flood peaks in these areas will increase with extreme storms (Sharma et al., 2018).

1.2 Climate Model Findings

In order to better understand the future of these trends, climate models are a useful tool. Global Climate Models (GCM's) can reveal more about the physical and dynamic processes of the earth surface-atmosphere system (Jiang et al., 2007). They were designed to simulate the response of the global climate system to greenhouse gas (GHG) forcings, and to guide emissions policies to mitigate climate change impacts (Kundzewicz & Stakhiv, 2010). These GCMs provide plausible patterns of climate change globally and can be useful for identifying projected climatological trends and examining the hydrological cycle. Global Climate Models (GCMs) are often downscaled to achieve more detailed and accurate regional information.

One set of GCMs that is widely used by the broad climate research community is the Coupled Model Intercomparison Project (CMIP), an ensemble group of models developed by The World Climate Research Programme's (WRCP) Working Group on Coupled Modelling (WGCM), with input from the International Geosphere–Biosphere Programme's (IGBP) Analysis, Integration and Modeling of the Earth System (AIMES) project (Taylor, Stouffer, & Meehl, 2012). The ensemble includes models developed by many different agencies around the world. The model experiments are integrated using atmosphere-ocean global climate models (AOGCMs). The models include an interactive representation of the atmosphere, ocean, land, and sea ice, and respond to various atmospheric constituents, such as greenhouse gases (Taylor et al., 2012). Different emissions scenarios are represented by Relative Concentration Pathways

(RCPs). The different RCPs utilize different radiative forcing scenarios to encompass a wide range of future anthropogenic activity, including a high-emissions scenario (RCP8.5), a midrange mitigation emissions scenario (RCP4.5), and a low emissions "peak-and-decline" scenario (RCP2.6) (Taylor, Stouffer, & Meehl, 2012). Scenarios RCP2.6, RCP4.5, and RCP8.5 have estimated radiative forcing of 2.6 Wm^{-2}, 4.5Wm^{-2} and 8.5Wm^{-2}, respectively, by 2100 (Kharin, Zwiers, Zhang, & Wehner, 2013).

There is still significant uncertainty remaining in GCMs. They are known to contain errors because of the simplification of climate representation, assumptions about climate processes, and potential errors in forcing data (Buytaert, Célleri, & Timbe, 2009). Climate sensitivity, heat uptake by the oceans, aerosol forcing, and GHG emissions are all still somewhat uncertain parameters (Dessai & Hulme, 2013). The utility of the CMIP data set lies in the ensemble of GCMs, as any single GCM can be greatly impacted by the modeling uncertainties and give unreliable results. The use of different GCMs can therefore produce largely different results (Jiang et al., 2007). Substantial biases also exist for GCMs, especially for the simulation of sub-daily precipitation extremes (Prein et al., 2017). These biases are generally related to convective rainfall, which climate models struggle to simulate accurately (Allan & Soden, 2008) due to the coarse spatial resolution and parameterization of convective processes.

However, the models have been found to agree relatively well with observational evidence in the extratropics (Kharin, Zwiers, Zhang, & Wehner, 2013). The 20 year precipitation event, for example, had discrepancies of +/- 20% (Kharin, Zwiers, Zhang, & Wehner, 2013). Model results for the tropics and subtropics, however, are still highly uncertain.

Studies employing the CMIP5 ensemble have found that changes in precipitation modeled by CMIP5 are more uneven than present day, with 1/5 of the projected daily increase occurring over just 2 days each year, and 70% in just 2 weeks (Pendergrass & Knutti, 2018). Unevenness increased with increased emissions scenarios, and suggests that unevenness at an hourly scale would be even greater (Pendergrass & Knutti, 2018). Kharin et al. (2013) found that relative increase in extreme precipitation intensity increased more than annual mean precipitation in the models. For the 20-year event, the multi-

model average increase ranged from about 6% to more than 20% across the different RCPs (Kharin et al., 2013).

Based on analysis employing climate models, today's 100-year flood may become more frequent for many watersheds in the future (Šraj et al., 2016). One study analyzed seasonal extreme precipitation across the US with CMIP5, and found that extreme precipitation intensified by 27% over the summer months (Prein et al., 2017). For the Kansas City area, the mean and moderate (75%) events indicated decreasing trends in summer, but the change signal for extreme events (99.95%) was increasing, with the exceedance probability increasing by about 50%. For the winter months, the mean, moderate, and extreme events were all increasing, with the exceedance probability increasing by 100-200%. Observed increases of precipitation extremes have been found to be larger than the models predict. Studies suggest that GCM simulations may actually underestimate extreme precipitation in the Northern Hemisphere, providing conservative outputs (Condon, Gangopadhyay, & Pruitt, 2015).

GCMs have revealed a distinct link between extreme precipitation and temperature (Pendergrass & Knutti, 2018). Extreme precipitation is driven by increasing moisture and circulation, so extreme events occur with changes in the moist adiabatic temperature lapse rate, upward velocity, and temperature (Pendergrass & Knutti, 2018). Allan & Soden (2008) demonstrated this link using satellite observations and model simulations based on natural climate variability.

1.3 Precipitation Scaling Rate

The Clausius-Clapeyron (C-C) relation is widely held to be the standard for estimating future precipitation increases. This follows the concept that a warmer atmosphere can hold more moisture, leading to more precipitation (Pfahl, O'Gorman, & Fischer, 2017). In the C-C relationship, extreme precipitation intensity is proportional to the atmospheric water vapor content, which increases exponentially with temperature at the C-C scaling rate, approximately 7%/C (Wang et al., 2017). This

relationship assumes that the nature of atmospheric circulation does not change considerably, which has been questioned (Lenderink & Van Meijgaard, 2008). In fact, many GCMs predict changes in the atmospheric circulation and related changes in precipitation (Lenderink & Van Meijgaard, 2008). Most events could increase at or above the rate of moisture increase (Pendergrass, 2018).

Recent studies show that the C-C relation is too simplistic (Pendergrass, 2018). Based on observational data, the relationship between local temperature and precipitation extremes deviates significantly from the C-C scaling (Wang et al., 2017). At the hourly and sub-hourly scale, precipitation is driven by convective rainfall. In this case, increased latent heat release may intensify the upward motions, causing an increased scaling rate. In some studies, observational evidence of a 'superCC' relationship has been found for hourly and sub-hourly extremes (Kendon, Blenkinsop, & Fowler, 2018). In these instances, extremes increase twice as fast with raising temperatures as the C-C scaling when the temperature is greater than 12 degrees C (Lenderink & Van Meijgaard, 2008).

Observations show that the scaling rate increases in intensity for the more extreme events (Lenderink & Van Meijgaard, 2008). The actual scaling rate has been found to be very dependent on the region, temperature, and moisture availability, so simply extrapolating from historical data is not accurate (Prein et al., 2017). Based on the CMIP5 models, the Midwest region was found to have a 3-6%/K scaling rate for annual maximum precipitation (Pfahl et al., 2017). One study even suggested that precipitation extremes outside the subtropics have a better scaling rate with mean precipitation than water vapor content (O'Gorman & Schneider, 2009).

1.4 Statistical Methods for Estimating Extreme Precipitation

A wide range of statistical methods have been used for the design of hydraulic structures to protect against the impact of extreme events. Starting with either precipitation or streamflow data, the data is fit to a distribution, and the resulting probability distribution functions (PDFs) are used to

determine event probabilities. Using a distribution allows the extrapolation of probabilities beyond the available data set. Gumbel stated, ‘‘Some authors have tried to introduce upper and lower limits to the discharges, even though it is doubtful that such limits exist’’ (Katz, Parlange, & Naveau, 1989). Applied distributions include lognormal (LN), Gamma, Log-Pearson Type III (LP3), and extreme value distributions (Salas, Obeysekera, & Vogel, 2018).

One of the most common methods is the Block Maxima (BM) approach, with the application of extreme value theory to Annual Maximum Series (AMS), consisting of the largest event for each year. Because the distributions of hydrologic variables are heavy tailed (Katz, Parlange, & Naveau, 1989), the maxima can be described by one of the three extreme value distributions: Gumble, Frechet and Weibull (Sarhadi & Soulis, 2017). More commonly, Generalized Extreme Value (GEV) incorporates all three (Sarhadi & Soulis, 2017). Various methods have been used for parameter estimation, most notably maximum likelihood and L-moments (Frei, Schöll, Fukutome, Schmidli, & Vidale, 2006; Katz, Parlange, & Naveau, 1989). Regional frequency analysis is often used to improve estimations at a given site using observations from other sites, if the region is relatively homogeneous (Frei et al., 2006; Katz, Parlange, & Naveau, 1989). Studies have found that regional analysis can produce more accurate quantile estimates than a single site alone (Katz, Parlange, & Naveau, 1989). The analysis can also use the method of least squares to relate site variables to physical characteristics (Katz, Parlange, & Naveau, 1989). Bayesian methodology is also frequently used to provide more robust estimations (Katz, Parlange, & Naveau 1989).

A more flexible, but more complex approach is the Peak-over-Threshold (POT) method. Rather than the AMS, this method uses the Partial Duration Series (PDS), which consists of all values greater than a chosen threshold. Because this method includes more extreme values, it should provide more robust information about the extreme upper tail, and the resulting parameter estimation and final distribution quantiles should be more accurate (Katz, Parlange, & Naveau, 1989). Additionally, using a

larger sample of extremes could result in decreased uncertainty (Salas et al., 2018). Mallakpour and Villarini (2015) found more trends with the POT approach than AMS.

The biggest challenge with using PDS is threshold selection. Both the threshold and resulting POT values are important from both physical and statistical points of view, yet there are no universally accepted standard guidelines (Salas et al., 2018). Methods range from modeling the threshold using quantile regression (Salas et al., 2018), to simply selecting a desired frequency (2 events per year) (Mallakpour & Villarini, 2015). Other methods include the graphical method (GM) and the optimal bias robust estimation (OBRE) method (Solari & Losada, 2012). The POT method generally involves two steps: a Poisson process describing the rate of exceedance (frequency), and the Generalized Pareto distribution describing the excess over the threshold (PDS) (Mondal & Mujumdar, 2015).

These traditional statistical methods are based on one large assumption: the stationarity of extreme events. These estimates are assuming that the fluctuation of hydrological systems has an unchanging variability, and that key statistics, such as mean and variance, don't change with time (Salas et al., 2018). Substantial evidence shows that this assumption is false, and in fact, "stationarity is dead" (Milly et al., 2008). Small changes in the center of a distribution can have significant impacts on the distribution tail, leaving both frequency and intensity sensitive to change (Condon et al., 2015). This standard approach has led to abundant underestimation in the probability of extreme events (Chandra, Saha, & Mujumdar, 2015). The underestimation of designs can lead to catastrophic damage, and already has in several cases (Chandra et al., 2015). Climate variability and anthropogenic influences need to be incorporated into hydrologic estimates, leading to a shift away from traditional, stationary methods toward methods incorporating nonstationary aspects and parameters.

While there is lacking consensus on the best way to achieve nonstationary estimates, studies have approached the challenge in a variety of ways. Generally, the detection of a trend is a critical first step that influences the nonstationary analysis results (Šraj et al., 2016). One common method is to involve parameters that are dependent on time (Šraj et al., 2016). Strupczewski et al. (2001) conducted an analysis

with two time-dependent parameters in the Gamma and lognormal PDFs. Singh et al. (2013) developed nonstationary GEV and lognormal distributions. Gül et al. (2014) varied the GEV parameters with a GEV conditional density estimation network. Delgado et al. (2010) varied the location and scale parameters with time in the GEV distribution for the Mekong River. Parey et al. (2007) used the GP distribution, by making both the events and parameters time-dependent. These time-varying parameters make the large assumption that parameters will continue to change in the future as they did in the past, which may not be the case (Šraj et al., 2016).

Another general method is to vary distribution parameters with relevant hydrological covariates, including climatological variables or land use (Šraj et al., 2016). These covariates range from the El-Nino Southern Oscillation index (ENSO) to local mean temperatures (Mondal & Mujumdar, 2015). Prosdocimi et al. (2015) used the POT method and modeled the resulting GEV parameters as functions of both time and other climatological variables. Delgado et al. (2014) applied the lognormal distribution with the scale parameter as a function of the non-stationary standard deviation of an atmospheric circulation index representing Western Pacific monsoon activity.

Some studies have incorporated climate model projections into their estimates as well. Gilroy and McCuen (2012) used GCM-estimated maximum annual 24-hour precipitation, as well as urbanization estimates, to determine flood frequencies. Condon et al. (2015) incorporated CMIP5 projections into nonstationary GEV analysis. Šraj et al. (2016) tested the GEV distribution for one or two parameters in different combinations, using maximum likelihood and Bayesian method through Markov Chain Monte Carlo simulations, and found that parameters dependent on annual precipitation and time worked best at their stations. Their results showed that event precipitation and annual precipitation are usually highly correlated. Significant differences between the non-stationary and stationary quantile estimates were evident, with the stationary 100-year event corresponding to a nonstationary 20-year event. Cheng and AghaKouchak (2014) found similar results when they compared stationary and nonstationary

Intensity-Duration-Frequency (IDF) curves in the US and suggested that government guidelines for IDF curves need to be updated.

Uncertainty always accompanies these statistical-based predictions of extreme events. The major source of uncertainty in IDF estimates is an insufficient quantity and quality of data and resulting parameter uncertainty (Chandra et al., 2015). The sample uncertainty also increases with increasing return period (Šraj et al., 2016). Additional uncertainty is introduced with GCMs, either from using multiple models, or from using only one (Chandra et al., 2015). They found that model uncertainty is small compared to parameter uncertainty.

1.5 Study Overview

It is vital for engineers to have improved, more robust methods in order to create resilient and sustainable designs within the context of an evolving hydrologic cycle. The purpose of this study is to characterize the projected changes in storm intensity, duration, and frequency for the Kansas City area to support engineering design. Historical gage data from the National Oceanic and Atmospheric Administration (NOAA) and projected data from the CMIP5 model ensemble are used. Precipitation trends in both the model and gages for the Kansas City area are analyzed. The relationship between PDS event frequency and annual precipitation are employed to connect climate model data (available as monthly averages) to the gage PDS data and develop adjusted PDS data sets. IDF curves were developed from this PDS data, using NOAA methods, to incorporate modelled precipitation data through the year 2100.

Chapter 2: Methods

2.1 Study Area

2.1.1 Geographic Setting

The Kansas City metropolitan area straddles the Eastern edge of Kansas and the western edge of Missouri. It includes 8 main counties including Johnson County, Miami County, Cass County, Jackson County, Ray County, Clay County, Platte County, and Leavenworth County and is home to Kansas City, KS and Kansas City, MO and many surrounding suburbs such as Overland Park (MARC, 2019). The major waterbodies include the Missouri River, the Kansas River, and the Blue River. All of the area's runoff eventually drains to the Missouri River, which flows from the Northwest through downtown Kansas City. The Kansas River flows from the West, flowing into the Missouri river in downtown Kansas City. The Blue River flows from the South, converging with the Missouri River northeast of Kansas City.

2.1.2 General Climate Conditions in Kansas City

The climate of the Kansas City area is highly seasonal. The hot season (June–September) generally has average daily high temperatures of 80-90^{o}F. The cold season (December– February) has temperatures of 40-45 ^{o}F. The area receives about 39 inches of rainfall annually, with the most precipitation occurring April through September and average monthly rainfall of 3-6 inches (USDC, NOAA, NESDIS, 2019). The average relative humidity ranges from about 62% (April) to 71% (December). The eastern part of Kansas receives very moist air from the Gulf of Mexico during the summer, which often brings air-mass thunderstorms.

2.2 Datasets

2.2.1 CMIP5 Data

Projected monthly precipitation data was retrieved from downscaled Global Climate Model (GCM) simulations from the Coupled Model Intercomparison Project Phase 5 (CMIP5) (WCRP, 2019). The collection of models in the CMIP5 project are run under the same historical and future scenarios relative to radiative forcing and other variables, making them ideal for comparison. We analyzed only the models with the full 'r1ip1' simulation with all three RCPs, which includes 20 of the ensemble CMIP5 members. The data sets consist of historical run data (1950-2005) and future run data (2005-2100). We analyzed mean daily precipitation rate (mm/day) at the monthly time scale.

The 1° x 1° (approximately 100 km) spatial resolution of the CMIP5 climate models is too coarse to represent the local hydrologic cycle. To overcome this scale issue, statistically downscaled climate data is used from the CMIP5 Climate and Hydrology Projections datasets (Bureau of Reclamation 2013). The monthly bias-correction spatial disaggregation (BCSD) downscaled version of CMIP5 was used for this study. The BCSD process involves first regridding the model data to a common 1^{o} resolution (referred to as "REGRID"). Bias-correction is then guided by gridded observational data sets. The spatial resolution of the downscaled data is 1/8^{o}, or 12km X 12km.

Grid cells were chosen at various spots across the KC metro area. The 20 CMIP5 models used are shown in Table 1 below. The selected geographic locations used are shown in Table 2 below. The climate model data was obtained from:

https://gdo-dcp.ucllnl.org/downscaled_cmip_projections/dcpInterface.html#Projections:%20Complete%20Archives

Table 1:CMIP5 Models used in this study and their agency.

Model Name	Agency
bcc-csm1-1	Beijing Climate Center, China Meteorological Administration, China
canesm2	Canadian Centre for Climate Modeling and Analysis, Canada
ccsm4	National Center of Atmospheric Research, USA
cesm1-cam5	National Center of Atmospheric Research, USA
csiro-mk3-6-0	Commonwealth Scientific and Industrial Research Organization/Queensland Climate Change Centre of Excellence, Australia
fgoals-g2	Institute of Atmospheric Physics (LASG) and Center for Earth System Sciences (CESS)
fio-esm	First Institute of Oceanography (FIO)
gfdl-cm3	NOAA Geophysical Fluid Dynamics Laboratory, USA
gfdl-esm2g	NOAA Geophysical Fluid Dynamics Laboratory, USA
gfdl-esm2m	NOAA Geophysical Fluid Dynamics Laboratory, USA
giss-e2-r	National Aeronautics and Space Administration (NASA), USA
hadgem2-ao	Met Office Hadley Center, UK
hadgem2-es	Met Office Hadley Center, UK
ipsl-cm5a-lr	Institut Pierre Simon Laplace, France
ipsl-cm5a-mr	Institut Pierre Simon Laplace, France
miroc5	Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology, Japan
miroc-esm	Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology, Japan
miroc-esm-chem	Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology, Japan
mpi-esm-lr	Max Planck Institute for Meteorology, Germany
mpi-esm-mr	Max Planck Institute for Meteorology, Germany
mri-cgcm3	Meteorological Research Institute, Japan
noresm1-m	Norwegian Climate Center, Norway
noresm1-me	Norwegian Climate Center, Norway

Table 2:Geographical locations of CMIP5 data

Location Number	Latitude	Longitude
1	39.06	-94.56
2	38.94	-94.69
3	38.81	-94.81
4	39.19	-94.56
5	38.94	-94.44

2.2.2 Gage Data

Rain gage observations were obtained from NOAA gage data at six locations (including one location with both an hourly and daily gage) across the Kansas City Metropolitan area, as shown in Figure 1 below (USDC, NOAA, NESDIS, 2019). In order to maintain consistency for comparison with CMIP5 data, only gages covering the full historical CMIP5 period (1950-2005) were selected. This includes five gages with daily data and two gages with hourly data. The precipitation data includes only rain events. Snow was assumed to be inconsequential for our purposes. The gages include Bonner Springs (Bonner), Downtown Airport (DTA), Olathe, Paola, Stilwell, and Unity Village (Unity).

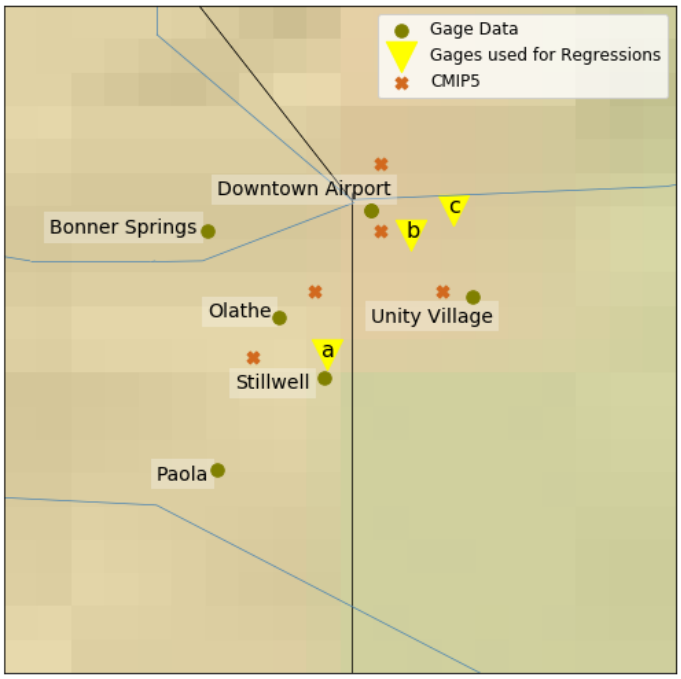

Figure 1: Rain Gage and CMIP5 Geographic Locations

Table 3: NOAA Gage Information

Gage Name	Gage ID	Frequency	Start	End	Latitude	Longitude
Bonner Springs	USC00140957	daily	1938	2018	39.06	-94.90
Downtown Airport	USW00013988	daily	1893	2017	39.10	-94.58
Olathe	USC00145972	daily	1893	2009	38.89	-94.76
Paola	USC00146209	daily	1895	2009	38.59	-94.88
Stillwell	USC00147809	daily	1939	2010	38.77	-94.67
Downtown Airport	COOP:233219	hourly	1950	2013	39.10	-94.58
Unity Village	COOP:234846	hourly	1949	2013	38.93	-94.38

Table 4: Additional NOAA Gages used for regressions

Gage Name	Gage ID	Frequency	Start	End	Latitude	Longitude
Stanley	COOP:147756	hourly	1982	2013	38.82	-94.67
Swope Park	COOP:234374	hourly	1948	1977	39.05	-94.50
Independence	USC00234158	daily	1973	2018	39.06	-94.39

Original data sets were constrained to fixed times once every day or hour. This was done to maintain consistency between all durations. The hourly gage data was aggregated to constrained 2-hour, 6-hour, and 12-hour increments (hours 0-12, etc.), and daily data was aggregated to constrained 2-day, 7-day, and 30-day increments, consistent with NOAA's methodology (NOAA, 2013). Constraining the data also prevented the reporting of a single event as multiple overlapping events. However, due to the constrained nature of the data, it is likely that the resulting AMS and PDS would not capture the true-interval maximums, resulting in an underestimation of storm frequencies. To account for this, correction factors were applied to the AMS and PDS data, as used by NOAA (Noaa, 2013). Correction factors are shown in Tables 5-6 below.

Table 5: Correction factors for daily AMS data due to constrained observations (NOAA, 2013).

Duration (days)	1	2	3	4	7	>7
Correction Factor	1.12	1.04	1.03	1..02	1.01	1

Table 6: Correction factors for hourly AMS data due to constrained observations (NOAA, 2013).

Duration (hours)	1	2	3	6	>6
Correction Factor	1.09	1.04	1.02	1..01	1

2.2.2.1 Missing Data

The raw data included both accumulated and missing precipitation data. NOAA procedure was followed for all accumulated data. The accumulated amount was evenly distributed across the entire accumulation period. Many of the gages had significant amounts of missing data (up to 5 consecutive

years in some cases). In order to more accurately test for and compare trends across the entire time period (1950-2005), missing gage data was filled in (as much as possible) using linear regression analysis.

A linear relationship was developed between the two variables, where y is the dependent variable (precipitation values for the gage with missing data) and x is the independent variable (precipitation values for the other gage). The best fit line with the minimum error was developed as a linear equation:

$$\hat{y} = b + a * x \qquad \textit{Equation 1}$$

Goodness of fit was determined based on the coefficient of determination (R^2), which is the proportion of the variation of the predictand (SST) accounted for by the regressions (SSR), described in equations 2-4 below (Wilks, 2011). R^2 values were calculated using the python 'stats.linregress(x,y)' function.

$$SST = \sum_{i=1}^{n} (y_i - \bar{y})^2 \qquad \textit{Equation 2}$$

$$SSR = \sum_{i=1}^{n} (\hat{y}(x_i) - \bar{y})^2 \qquad \textit{Equation 3}$$

$$R^2 = \frac{SSR}{SST} \qquad \textit{Equation 4}$$

To achieve better fit, regression analysis was performed seasonally, where winter is December-February (DJF), spring is March-May (MAM), summer is June-August (JJA), and fall is September-November (SON).

For each daily gage with missing data, the other five daily gages were tested, as well as all other nearby NOAA gages with overlapping time periods. Any missing data that could not be infilled was classified as 'nan' and excluded from analysis. For the hourly gages, days with any number of hours

missing were classified as missing data. Hourly gages were filled in both at the hourly and daily timescale. Regressions were fit at both the log and non-log scale. While different variables and methods of regression were tested, simple linear regression with nearby gage precipitation data provided the best fit in all cases. The regression with the highest possible R^2 value was chosen for each gage. Due to the variable time-periods and missing data, multiple regressions were required for each gage. Not all seasons were filled in with every gage, depending on R^2 value. A threshold of $R^2 > 0.2$ was used. Missing data was filled as much as possible, but some missing data remained. This was assumed to be negligible for daily data in further analysis. More complete data was available at the daily timescale than hourly. At the hourly timescale, insufficient data was available, and missing data was left as 'nan'. An example regression fit is shown in Figure 2 below. Table 7 below lists the gage names used to infill each of the six gages. More regression information can be found in Tables 11-16 in the appendix.

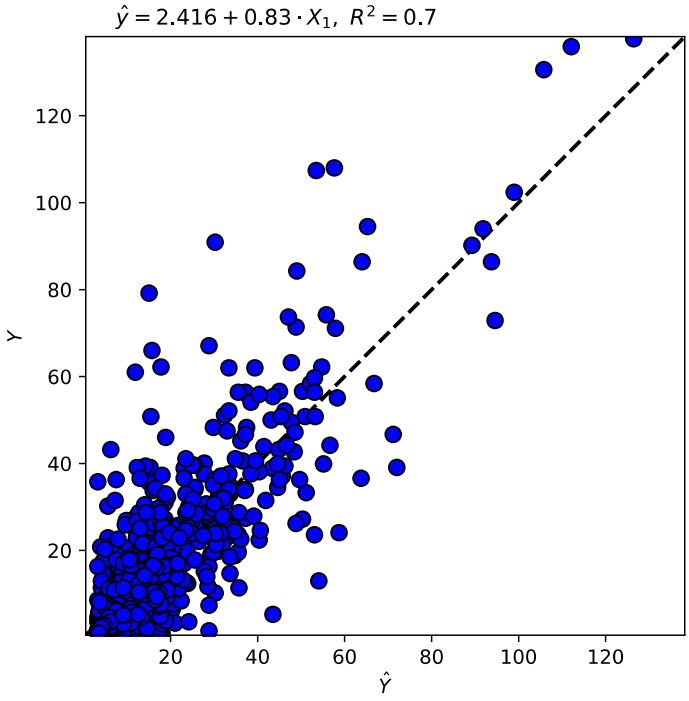

Figure 2: Stillwell (x) Winter (DJF) regression for Olathe (y)

.

Table 7: Linear Regression Information, where Regression #1 is used first to infill as much as possible, then Regression #2, and so on.

Gage y	**Gage X**			
	Regression #1	**Regression #2**	**Regression #3**	**Regression #4**
Unity	Stanley	Swope Park	Downtown Airport	
Stilwell	Olathe	Paola		
Olathe	Stilwell	Bonner Springs	Paola	Downtown Airport
Bonner	Olathe	Independence		
Paola	Stilwell	Olathe	Downtown Airport	
DTA	Gladstone	International Airport	Olathe	

2.3 Bias Correction

Before comparing the model data to gage data, the two were analyzed for any model biases. To do this, we compared the model's historical data (1950-2005) to the gage data during the same time period. Empirical Cumulative Distribution Functions (CDFs) were developed for each gage. First, an average was taken across all the models to compare to the gage (Figure 3a). However, the effect of taking an average simply minimized the variance of the model values and proved an ineffective method for comparison. To overcome this limitation, individual model ensemble members are considered instead of

the model-averaged CDF. An example for the noresm1-me model shows the difference between the model average (Figure 3a) and an independent model (Figure 3b).

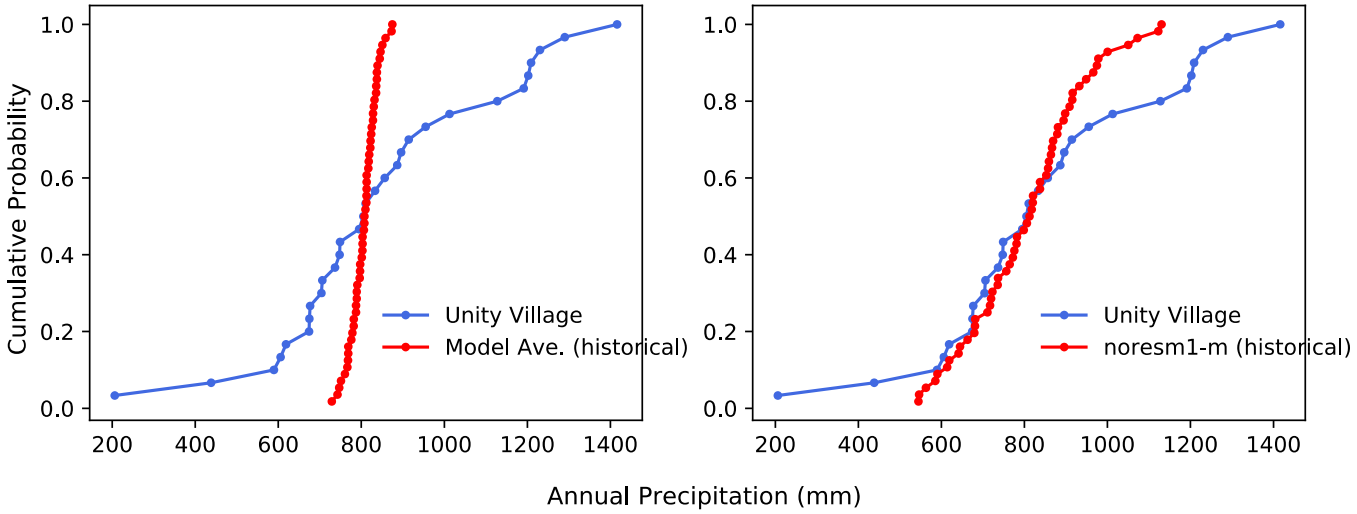

Figure 3: Empirical CDFs for Unity Village and historical model data. The model ensemble average is shown on the left (a) and an example of an individual model ensemble member (noresm1-m) is shown on the right (b).

As shown above, comparing individual historical model ensemble members to each gage provides a more representative comparison between the model and the gage data. While some models aligned well with the gages, many showed significant biases. Many of the models underestimate the tail, but the overall bias was not consistent across all 20 models and 6 gages. This inconsistency can be seen in Figures 4a-d below. Unity Village is represented well by the 'miroc-esm-chem' model in the middle section, but it underestimates the tail (Figure 4a). The 'mpi-esm-mr'model underestimates Paola throughout (Figure 4b). Unity Village is underestimated throughout by the 'ccsm4', with significant overestimation in the lower tail (Figure 4c). Olathe is replicated decently well by the 'gfdl-cm3' model, but the upper tail is once again underestimated (Figure 4d). Overall, there was significant bias across the models.

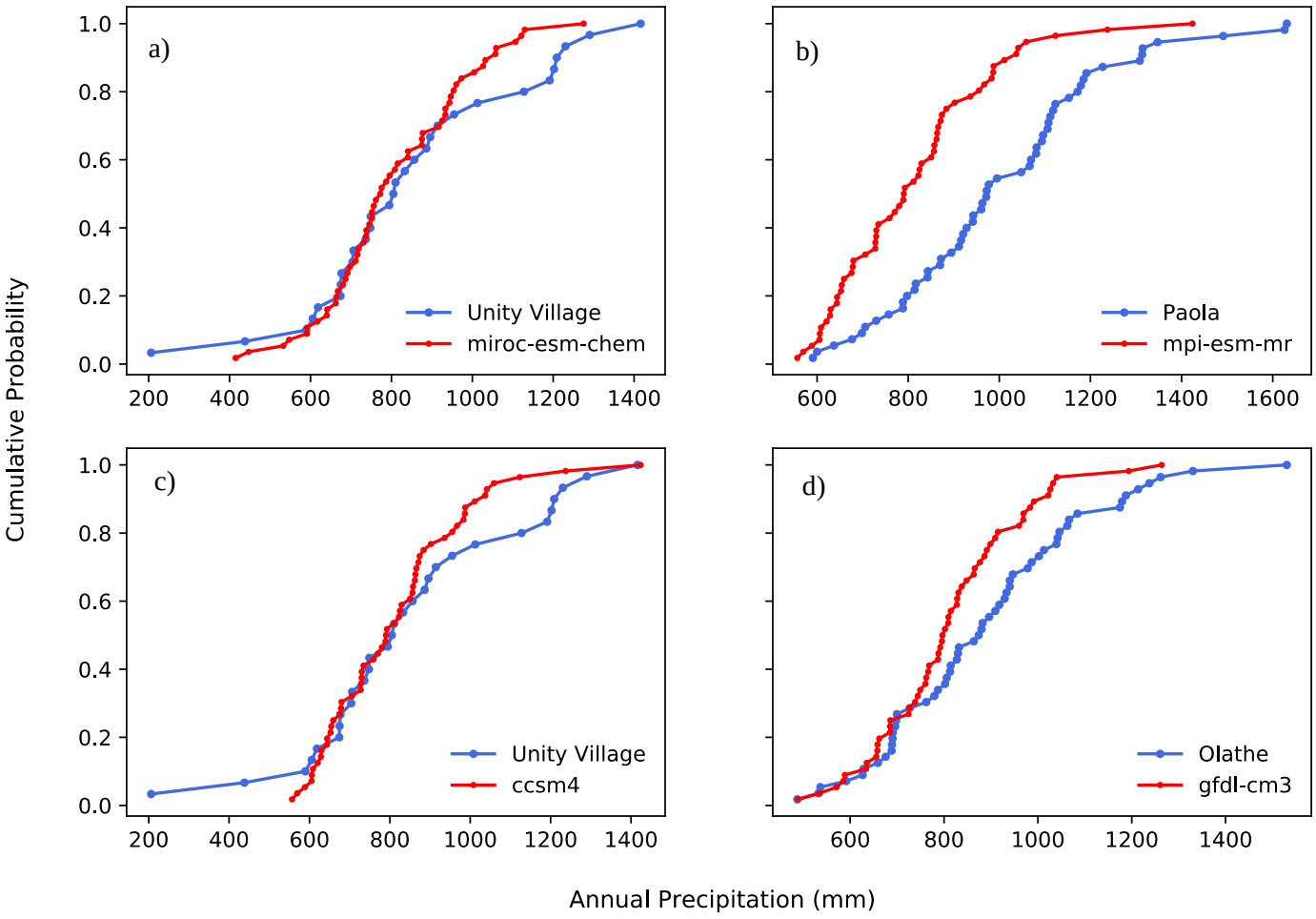

Figure 4: Comparison of empirical CDFs for annual precipitation from gage observations and model ensemble members. The data shown is for a) Unity Village and 'miroc-esm-chem', b) Paola and 'mpi-esm-mr', c) Unity Village and 'ccsm4', and d) Olathe and 'gfdl-cm3'.

Many different methods have been used to address model bias and get more relevant information from climate models. A relatively simple method is CDF matching (CDFm), which has been successfully used in many studies (Li et al., 2010). This approach maps the distribution of model data onto that of an observed data set. For annual precipitation (x), and both the model (m) and observations (o), the following equation is used (Li et al., 2010):

$$x_{m-p.adjust} = F_{o-c}^{-1}\left(F_{m-c}(x_{m-p})\right) \qquad \textit{Equation 5}$$

Where *F* is the CDFs. A limitation of this method is its inability to incorporate nonstationary aspects.

We employed this method to adjust the future CDF for each ensemble member based on the bias in the historical model. First, we fit distributions to the gage annual mean precipitation data and historical model annual precipitation data. The following distributions were tested: Normal, Lognormal, Extreme Value (I), Pearson Type III, and Log-Pearson Type III. The distribution with the best fit for each data set was used. Examples of the resulting CDFs for select models are shown in Figure 5 below.

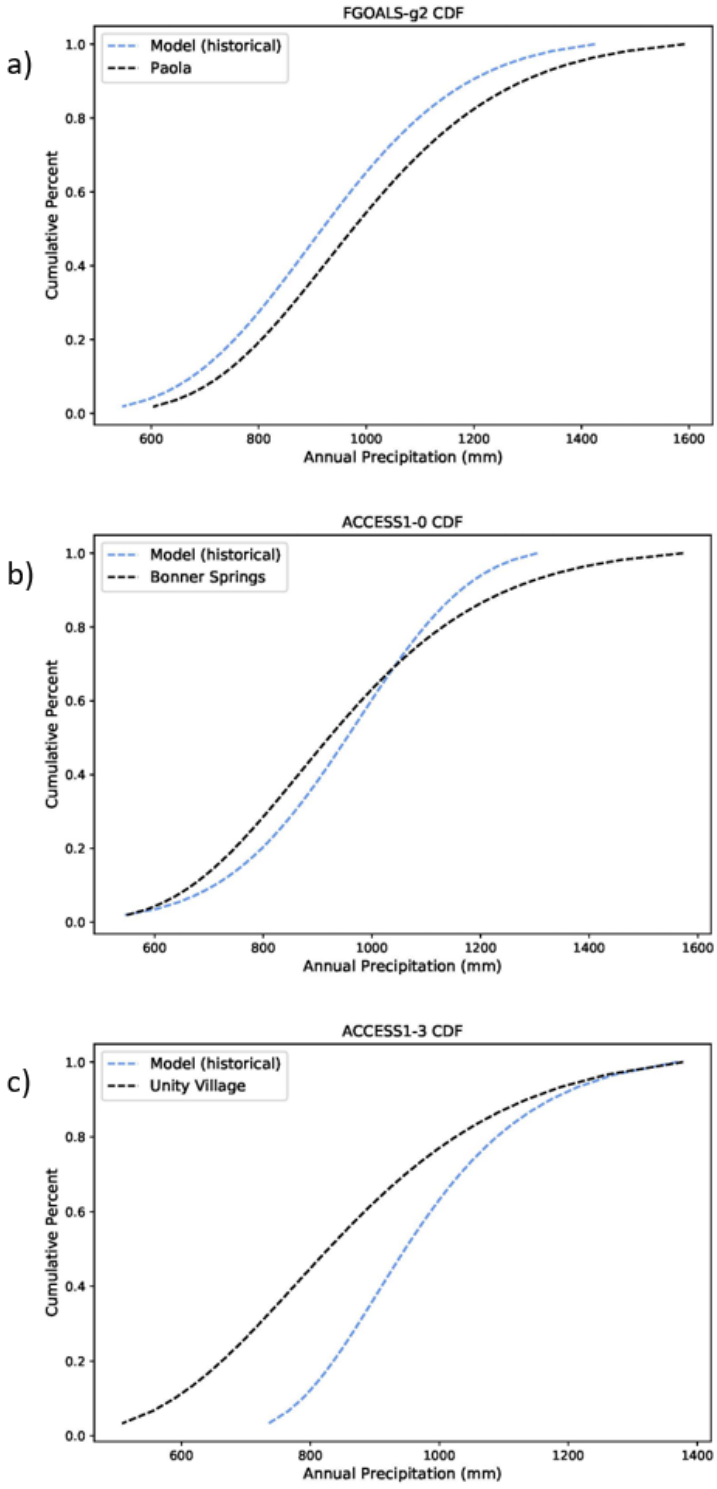

Figure 5: Fitted CDFs for annual precipitation where the black line is CDF for the observed precipitation, and the blue line is the fit CDF for the historical model. The data shown is for a) Paola and 'FGOALS-g2', b) Bonner Springs and ACCESS1-0, and c) Unity Village and ACCESS1-3.

These CDFs were then used to bias-correct the future model's annual precipitation data. Each value in the future model's empirical CDF was assigned a probability based on the historical model distribution. This probability was entered into the gage distribution to get the final value. The resulting values make up the final (adjusted) CDF for future annual precipitation. This was done for all combinations of the 20 models and six gages. As seen in Figure 6 below, this prevents significant underestimation of future values due to model bias. This also had a significant impact on the overall trend for annual precipitation, providing more accurate representation. If there is underestimation due to model bias, the jump from historical gage data (through 2005) to the model projections (starting in 2006) could cause an underestimation in the significance of an increasing trend, or even identify a negative trend that does not exist (and vice versa for overestimation).

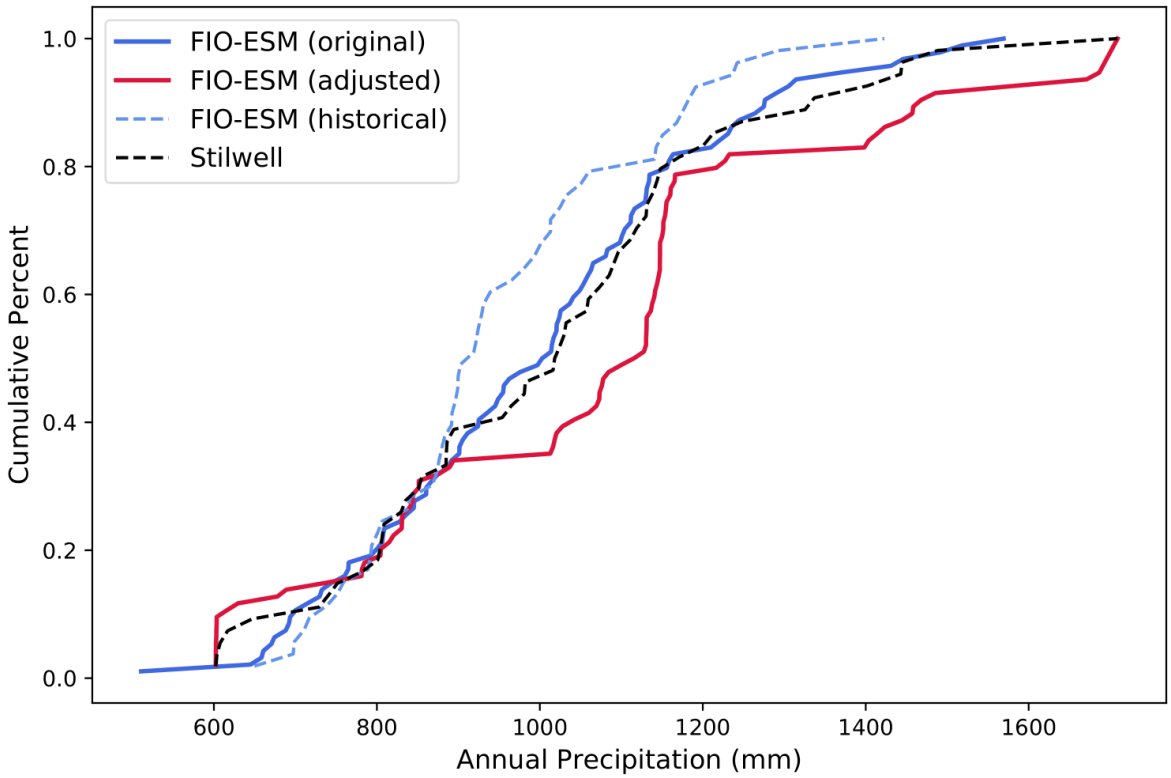

Figure 6: Annual Precipitation CDFs showing the bias correction of projected model precipitation for the FIO-ESM model at the Stilwell gage. The corrected data will provide a more representative assessment of future changes given by climate models.

2.4 Trend Analysis

The first goal of the analysis is to determine the presence of historical and projected precipitation trends. For this study, we used the Mann Kendall trend test to test for the significance of trends in precipitation data. Mann-Kendall is a non-parametric rank-based test that determines the direction and significance of a trend, without making underlying assumptions about the data distribution (Wilks, 2011). The null hypothesis (no trend) is rejected if the test statistic is larger than the critical value. The test statistic S is calculated as:

$$sgn(\theta) = \begin{cases} 1, if\ \theta > 0 \\ 0, if\ \theta = 0 \\ -1, if\ \theta < 0 \end{cases} \qquad \textit{Equation 6}$$

where n is the number of data points, x_j and x_k are the x values of time series *i* and *j*. The variance of the test statistic S is given by:

$$Var(S) = \frac{1}{18} * \left[n(n-1)(2*n+5) - \sum_t t(t-1)(2*t+5) \right] \qquad \textit{Equation 7}$$

Where n is the number of data points, and t is the number of ties. The test statistic for significance (z) is given by:

$$z = \begin{cases} \frac{X-1}{Var(S)^{1/2}}, if\ S > 0 \\ 0, if\ S = 0 \\ \frac{S+1}{Var(S)^{1/2}}, if\ S < 0 \end{cases} \qquad \textit{Equation 8}$$

The null hypothesis of no trend is rejected if the test statistic z is larger than the critical value $z_{\alpha/2}$. The Mann Kendall trend test was used to test for trends in annual precipitation, monthly

precipitation, precipitation events of different durations, and Partial Duration Series data. In addition to the Mann Kendall trend test, basic statistics, such as means, minimums, and maximums were also compared to further analyze the change in precipitation in the climate model data based on the historical model data (1950-2005) and the projected model data (2005-2100).

2.5 Frequency Analysis Methods

The NOAA Atlas 14 methodology (NOAA, 2013) was used to as the foundation for performing frequency analysis in this work. NOAA based all of their frequency analysis on Annual Maximum Series (AMS). They developed AMS by extracting the highest precipitation amount of each calendar year. They then developed PDS from AMS data using a simple conversion formula. We chose to develop PDS directly from the data, in order to better capture trends in the frequency of extreme events. NOAA also employed regionalization to develop more robust datasets. For simplicity, we only analyzed individual gage data. The remainder of the NOAA methodology was followed and is described below.

2.5.1 Annual Maximum Series

All original data (not the filled in data) is used to develop the AMS and PDS. Annual Maximum Series (AMS) data sets were first developed with the annual maximum rainfall for each duration. The procedure developed by NOAA was followed using specified criteria designed to only use reasonable maxima and exclude years with too much accumulated or missing data, especially occurring during the critical precipitation months or the "wet season". The Kansas City area lies in the "South Plains" region, where the "wet season" is defined as April-October for daily durations, and May-September for hourly durations.

2.5.1.1 AMS Acceptance Criteria

The first criterion is that less than 20% of total data is missing. No more than 20% of the wet season may be missing as well. Some of the data sets had periods where observations were not taken regularly, so the accumulated amounts were recorded over extended periods of time (NOAA, 2013). No more than 33% of the data can be accumulated, unless at least 66% of accumulated periods are less than the threshold accumulation period lengths (D_{thresh}). For durations less than 2 days, D_{thresh} is defined as matching the selected duration. For durations between 2 and 20 days, D_{thresh} is equal to half of the selected duration. For durations equal to or longer than 15 days, D_{thresh} is equal to 15 days. Less than 15% of wet season data can be accumulated, unless at least 66% of accumulated periods are less than D_{thresh}. Annual maxima data is rejected if it does not pass all these criteria, unless it is larger than 95% of the AMS data. The final AMS consists of all years passing the criteria.

2.5.1.2 Peak-Over-Threshold Method

Partial Duration Series (PDS) were then obtained using the peak-over-threshold (POT) method, where the PDS is composed of all values equal to or greater than the chosen threshold value. Threshold selection is extremely important, as choosing a threshold that is too high or too low can result in the overestimation or underestimation of recurrence intervals. Many different methods have been used for the selection of threshold values, with much debate as to which method is best (Salas, Obeysekera, & Vogel, 2018). One of the most common methods involves describing the rate of exceedance using a one-dimensional Poisson process, which has been shown to work reasonably well (Mondal & Mujumdar, 2015; Salas, Obeysekera, & Vogel, 2018). As a starting point, the threshold is set as minimum AMS value, as described by the US Water Resources Council (U.S. Water Resources Council, 1982).

Poisson Distribution

The number of exceedances of the threshold in t years, N(t), was assumed to follow a Poisson Distribution (Rosbjerg & Madsen, 1992). The probability mass function for Poisson, where $k \geq 0$ and μ is the shape parameter, is:

$$f(k) = \exp(-\mu)\frac{\mu^k}{k!} \qquad \textit{Equation 9}$$

To fit our exceedances to the Poisson distribution, the python function 'Scipy.stats.poisson(k, μ, loc)' was used, where μ is the mean and loc=0. This resulted in good fits, with R^2 values greater than 0.85.

The exceedance rate was then tested for each PDS, using the method described by Stedinger et al. (1993, p 18.38). They define the arrival rate as the ratio $\lambda = N/n$, where $\lambda \geq 1$, N is the number of selected events, and n is the total number of years. Assuming a Poisson model for frequency, a Partial Duration Series with $\lambda > 1.65$ should produce more reliable estimates (Stedinger et al., 1993). The arrival rate (λ) was calculated to be greater than 1.65 for each Partial Duration Series.

Grubbs-Beck Statistical Test

Each PDS was then tested for outliers, using the Grubbs-Beck statistical test. Both high and low outliers were defined as PDS values that depart significantly from the trend of the corresponding PDS values. Grubb's-Beck tests the hypothesis (H_0) that there are no outliers in the data set (Grubbs & Beck, 2019). The test statistic for a minimum outlier is:

$$G = \frac{\bar{Y} - Y_{min}}{s} \qquad \textit{Equation 10}$$

The test statistic for a maximum outlier is:

$$G = \frac{Y_{max} - \bar{Y}}{s} \qquad \textit{Equation 11}$$

Where:

$$\bar{Y} = sample\ mean$$

$$Y_{min} = minimum\ value$$

$$Y_{max} = maximum\ value$$

$$s = standard\ deviation$$

And the hypothesis of no outliers is rejected at a significance level α if:

$$G > \frac{N-1}{\sqrt{N}} \sqrt{\frac{t^2_{\alpha/N,N-2}}{N-2+t^2_{\alpha/N,N-2}}} \qquad \textit{Equation 12}$$

Where $t^2_{\alpha/N,N-2}$ is the upper critical value of the t-distribution with N-2 degrees of freedom and a significance level of α/N.

For this test, a significance level of $\alpha = 0.05$ is used. Low outliers were generally already eliminated through the selection criteria based on missing and accumulated data. Any high outliers were carefully investigated to determine their accuracy, and either corrected or removed.

2.5.2 Development of Intensity-Duration-Frequency Estimates

Intensity-Duration-Frequency (IDF) relationships were computed at individual stations using a frequency analysis approach based on L-moment statistics, as used by NOAA (NOAA, 2013). NOAA tested the fit of multiple distributions to AMS data, including the following: 3-parameter Generalized Extreme Value (GEV), Generalized Normal, Generalized Pareto, Generalized Logistic and Pearson Type III distributions; 4-parameter Kappa distribution; and 5-parameter Wakeby distribution. Based on their analysis, they selected the GEV distribution for all stations and durations, which provided an acceptable fit more frequently than any other distribution. The GEV distribution is also commonly used for the analysis of extreme events.

The Partial Duration Series (PDS) method generally assumes exponentially distributed peak exceedances. While the GEV distribution sometimes has a good fit for PDS data, Generalized Pareto (GP) distributions are often fit with more success (Rosbjerg & Madsen, 1992). In this study, some sites and durations were much better fit by GEV distributions, and some were better fit to GP distributions. The distribution with the best fit for each case was used based on the R^2 value, for each individual site and duration.

2.5.2.1 Generalized Extreme Value (GEV) Distribution

L-moment statistics were used to calculate the three parameters for the GEV distribution: shape (k), location (ξ), and scale (α) (Hosking & Wallis, 1997). The first four L-moments were calculated, as L1, L2, L3, and L4. First the ratios t3 and t4 are calculated:

$$t3 = \frac{L3}{L2} \qquad \textit{Equation 13}$$

$$t4 = \frac{L4}{L2}$$ *Equation 14*

These ratios can then be used to estimate k:

$$z = \frac{2}{3 + t3} - \frac{\ln(2)}{\ln(3)}$$ *Equation 15*

$$k = 7.859 * z + 2.9554 * z^2$$ *Equation 16*

The gamma function, or the generalized factorial, is used to estimate the last two parameters. The gamma function is:

$$\Gamma(n) = \int_0^\infty x^{n-1} e^{-n} dx$$ *Equation 17*

where:

$$n = k + 1$$ *Equation 18*

The last two parameters can then be estimated:

$$\alpha = \frac{(L2 * k)}{(1 - 2^{-k})^\Gamma}$$ *Equation 19*

$$\xi = L1 - \frac{\alpha * (1 - \Gamma)}{k}$$ *Equation 20*

The GEV distribution is then described with the following equations:

$$f(x) = \alpha^{-1} e^{-(1-k)y - e^{-y}} \qquad \textit{Equation 21}$$

$$y = \begin{cases} -k^{-1}\log(1 - k(x-\xi)/\alpha & k \neq 0 \\ (x-\xi)/\alpha & k = 0 \end{cases} \qquad \textit{Equation 22}$$

$$F(x) = e^{-e^{-y}} \qquad \textit{Equation 23}$$

$$x(F) = \begin{cases} \xi + \alpha\{1 - (-\log(F)^k\}/k & k \neq 0 \\ \xi - \alpha * log\{-\log(F)\} & k = 0 \end{cases} \qquad \textit{Equation 24}$$

To calculate storm intensity for given recurrence intervals, we used the python function 'scipy.stats.genextreme'.

2.5.2.2 Generalized Pareto Distribution

The POT method for PDS employs a Generalized Pareto (GP) Distribution (Katz, Parlange, & Naveau, 1989). The two GP parameters were estimated using the method of moments (MOM) (Rosbjerg & Madsen, 1992). Scale (α) and shape (k) are estimated from sample mean (μ) and variance (σ^2) using the following equations:

$$\alpha = \frac{1}{2}\mu\left(\frac{\mu^2}{\sigma^2} + 1\right) \qquad \textit{Equation 25}$$

$$k = \frac{1}{2}\left(\frac{\mu^2}{\sigma^2} - 1\right) \qquad \textit{Equation 26}$$

The probability density function is:

$$f(x) = \frac{1}{\alpha}\left(1 - k\frac{x}{\alpha}\right)^{1/k - 1} \qquad k \neq 0 \qquad \textit{Equation 27}$$

$$f(x) = \frac{1}{\alpha}e^{-x/\alpha} \qquad k = 0 \qquad \textit{Equation 28}$$

Estimates for α and k were produced using equations 25 and 26. These estimates were then used as first guesses in the python function “scipy.stats.genpareto.fit” to achieve better estimates.

Chapter 3: Results

3.1 Trend Analysis

Various methods were used to detect changes in both the gage data and the CMIP5 climate models. The trends in annual precipitation were first analyzed using the Mann Kendall trend test, where the associated confidence level is 1-p (if p=0.05, there is a 95% confidence level). The Mann Kendall trend test results for each gage annual precipitation (1950-2005) all showed increasing trends, however, only Stilwell and Unity Village were significant at a 95% confidence level. Table 8 shows the Mann Kendall trend test statistical results for all six gages.

Table 8: Mann Kendall Trend test 'p' and 'z' values for gage annual precipitation.

Gage	z	p
Paola	1.46	0.14
Bonner Springs	1.23	0.22
Downtown Airport	0.29	0.77
Olathe	0.29	0.77
Stilwell	2.30	0.02
Unity Village	2.05	0.04

There was a wide range of results between the different climate model ensemble members, however, more models had significant increasing trends in annual precipitation, based on the Mann Kendall results over the historical time period (1950-2005) and projected period (2006-2100). Figure 7 shows the significance levels (p-value) for models with increasing trends (Fig. 9a) and models with decreasing trends (Fig. 9b). When averaged across the models, the 'p' and 'z' values were similar for positive and negative trends (Tables 9-10). However, there were more models with increasing trends, and more were determined to be significant. The strongest increasing trends were found for RCP8.5. Overall, positive trends were identified in annual precipitation for both the models and gages.

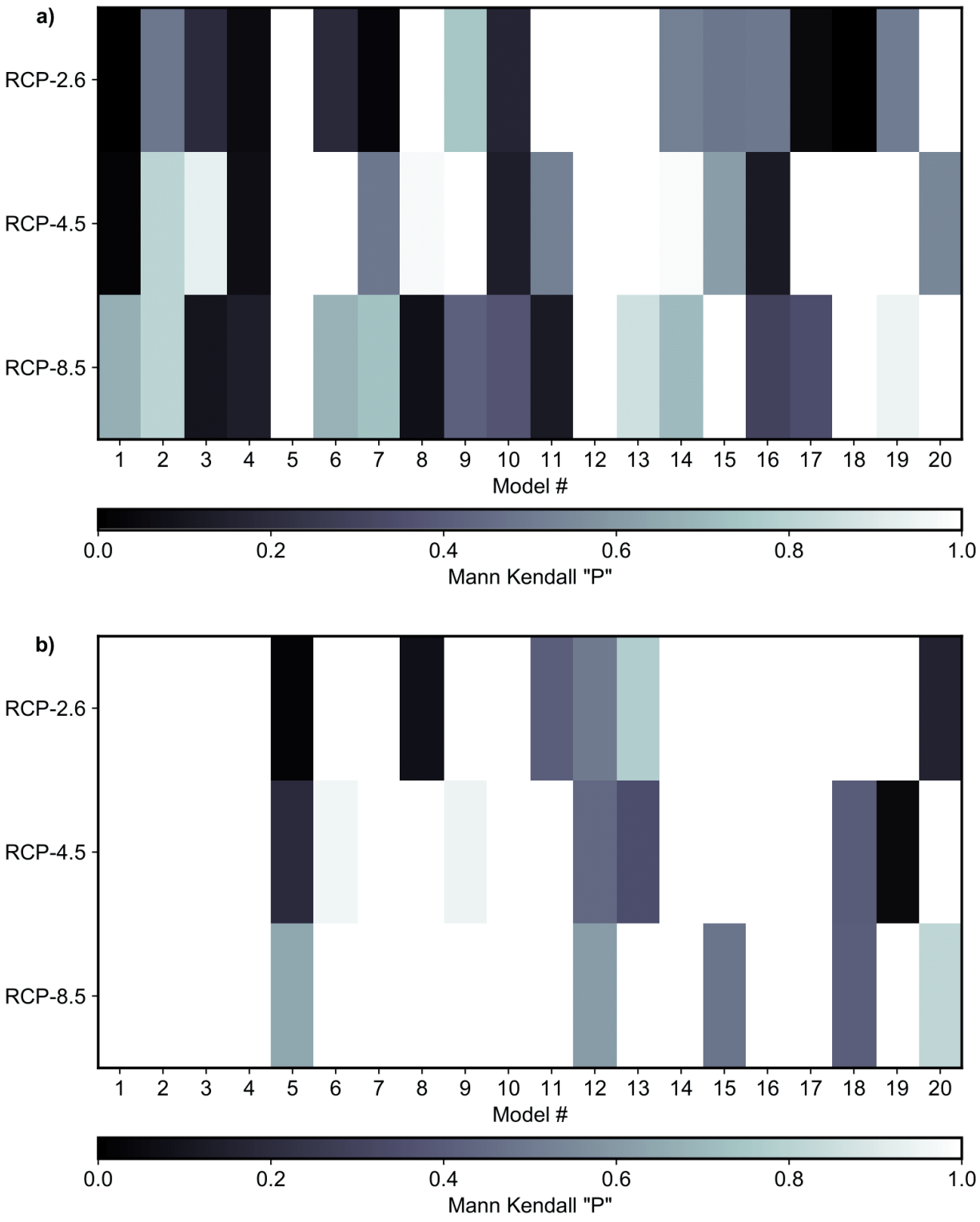

Figure 7: Significance of the Mann Kendall Trend test as given by the 'p' values for annual model precipitation, for a) positive trends and b) negative trends.

Table 9: Mann Kendall trend test 'p' and 'z' model results for annual precipitation with positive trends.

Model	'p' average	'z' average	Number of models with positive trends:		
			All	p < 0.01	p<0.05
Historical	0.53	0.07	15	0	0
RCP26	0.49	0.06	16	2	0
RCP45	0.52	0.06	13	2	1
RCP85	0.28	0.10	15	5	4

Table 10: Mann Kendall trend test 'p' and 'z' model results for annual precipitation with positive trends.

Model	p average	z average	Number of models with negative trends:		
			All	p < 0.01	p<0.05
Historical	0.53	-0.06	6	0	0
RCP26	0.59	-0.04	5	0	0
RCP45	0.48	-0.05	7	1	1
RCP85	0.33	-0.08	6	2	1

3.1.1 Seasonal Trends

Seasonal changes in both the model projections and gages were also explored. The strongest increasing signals were found in the winter (Dec-Feb) and spring (Mar-April) months, with slight decreases from June to August or September. The statistical change in the model from the historical (1950-2005) to future (2006-2100) time periods is first explored. For the CMIP5 model ensemble average, average monthly precipitation increased by up to 20% in March from the historical to future time periods (Figure 8). There was variability among the model ensemble members, with a range of about 40% for the percent change in monthly average (Figure 9). The model ensemble average for the maximum monthly precipitation for the historical (1950-2005) and future (2006-2100) time periods also showed increasing trends, with a more exaggerated increase overall (Figure 10). This indicates that climate models predict a seasonal change in the amount of precipitation with wetter springs and falls with dryer summers. However, the change in monthly maximum did not necessarily coincide with the monthly

average change. For example, while March had the largest increase in average monthly precipitation, there was very little change in maximum monthly precipitation. Also, average October precipitation increased by only about 5%, while maximum October precipitation had the largest increase of about 34%. Overall, these results suggest increased variability in precipitation in summer time with a decrease in the overall precipitation and an increase in the maximum monthly precipitation. In contrast, spring, fall and winter will see an increase in maximum precipitation and an increase in overall precipitation.

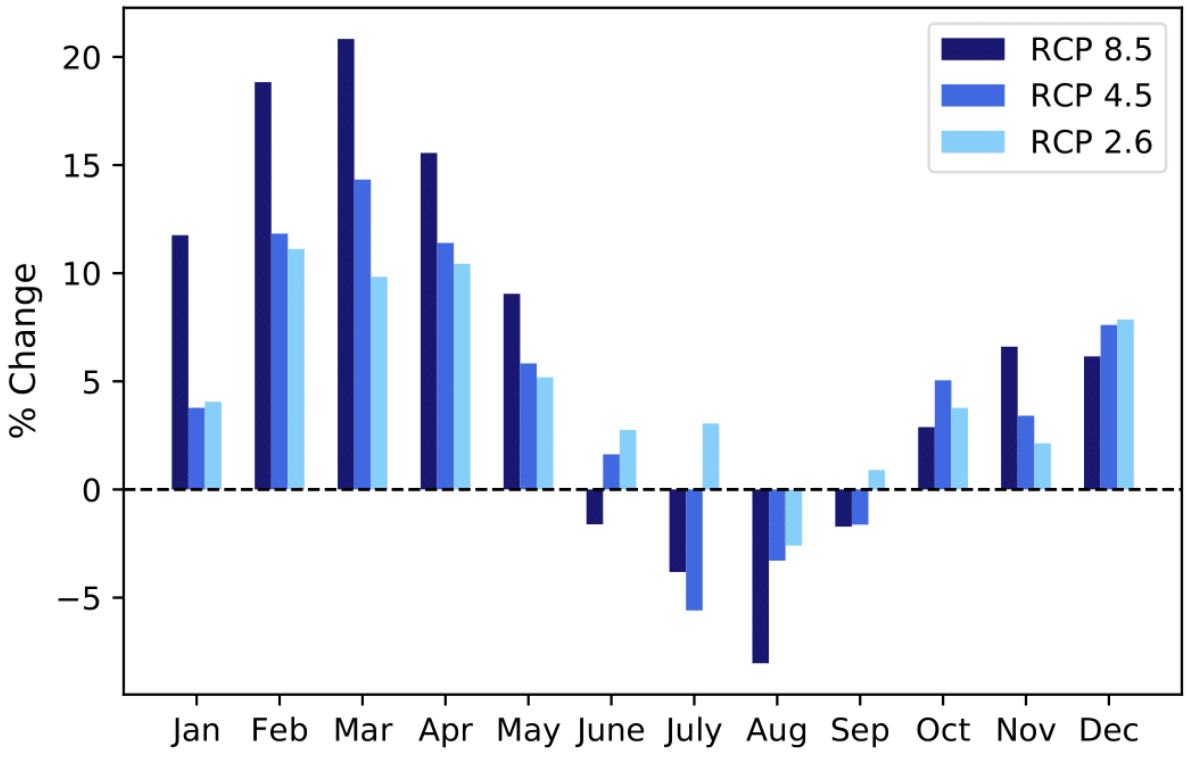

Figure 8: The average (across all models) change in average monthly precipitation from historical (1950-2005) to future (2005-2100) time periods.

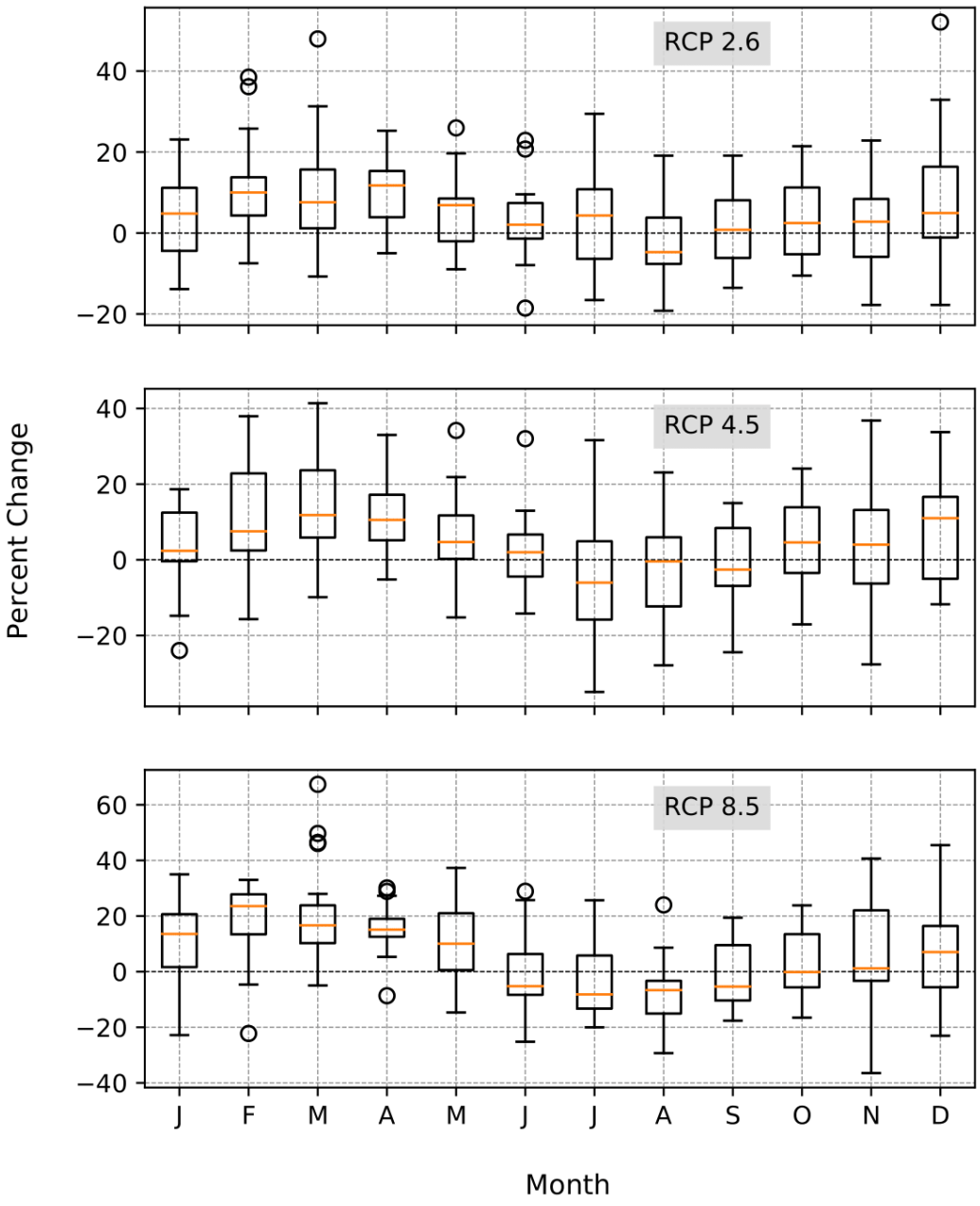

Figure 9: Box-and-whiskers plots showing the range across the 20 models. The circles are values determined to be outliers.

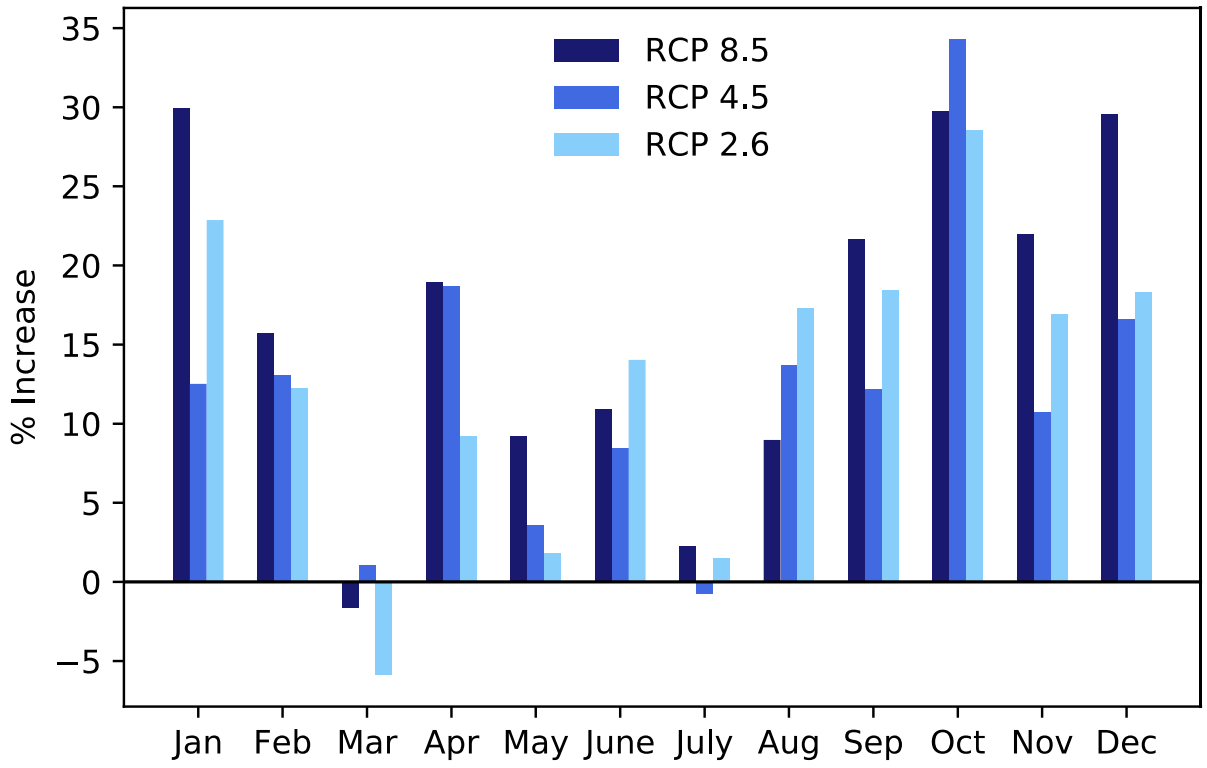

Figure 10: The average (across all models) change in maximum monthly precipitation from historical (1950-2005) to future (2005-2100) time periods.

The Mann Kendall Trend test was then used to test for significance monthly trends in the data. The results reflected the same general results. Average monthly precipitation was tested across the models and is shown in Figures 11 and 12 by the number of models with significant trends based on 95% and 90% confidence respectively. Overall, more models showed increasing trends in the winter and spring months, and decreasing trends in the summer and fall months (Figure 11). March and April had the most increasing trends, with 11 models (out of 20) with increasing trends in April with 90% confidence ($p < 0.1$), and 4 models with increasing trends in March with 95% confidence. July and October had the most models with significant decreasing trends, with 5 models having 90% confidence for both months, and 4 models having 95% confidence for October. The decreasing trends in October indicates a difference from what was seen by looking at the percent change in average precipitation (Figure 8). This is likely due to

higher magnitude of change in the models showing an increasing trend relative to the models showing a decreasing trend in October and indicates a large amount of uncertainty in the models. In contrast, the models show relative agreement for increasing trends during February, March, and April.

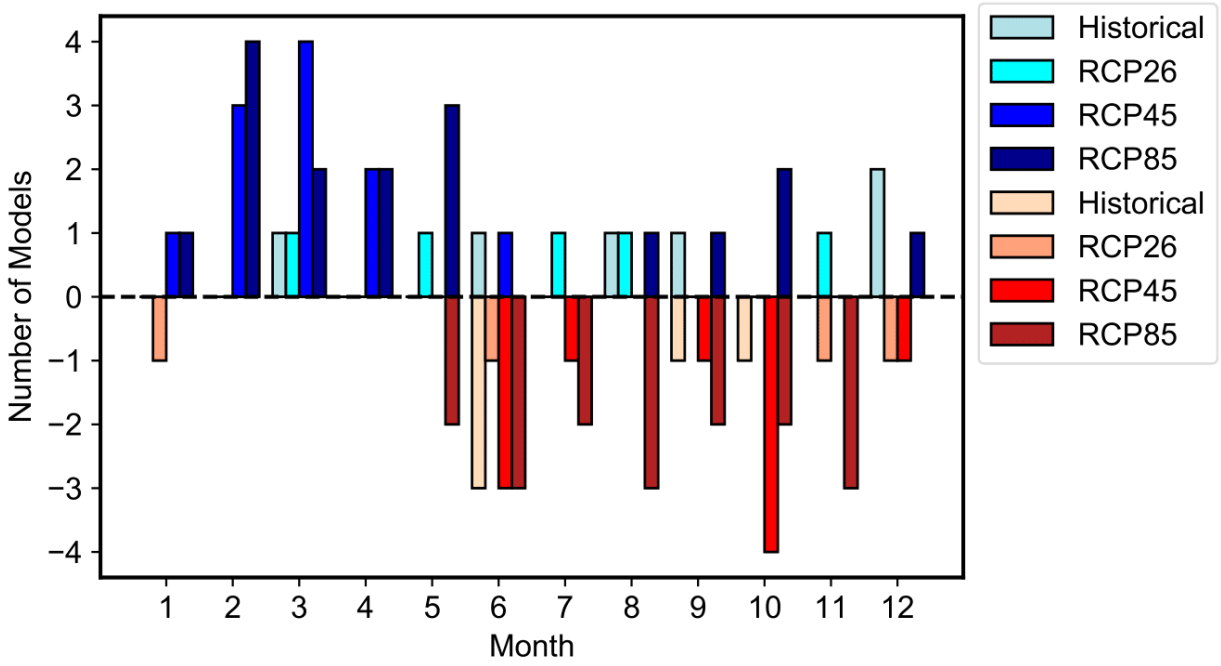

Figure 11: The number of models with Mann Kendall significance levels < 0.05 for monthly precipitation. Positive values represent models with significant positive monthly trends, and negative values represent models with significant negative trends.

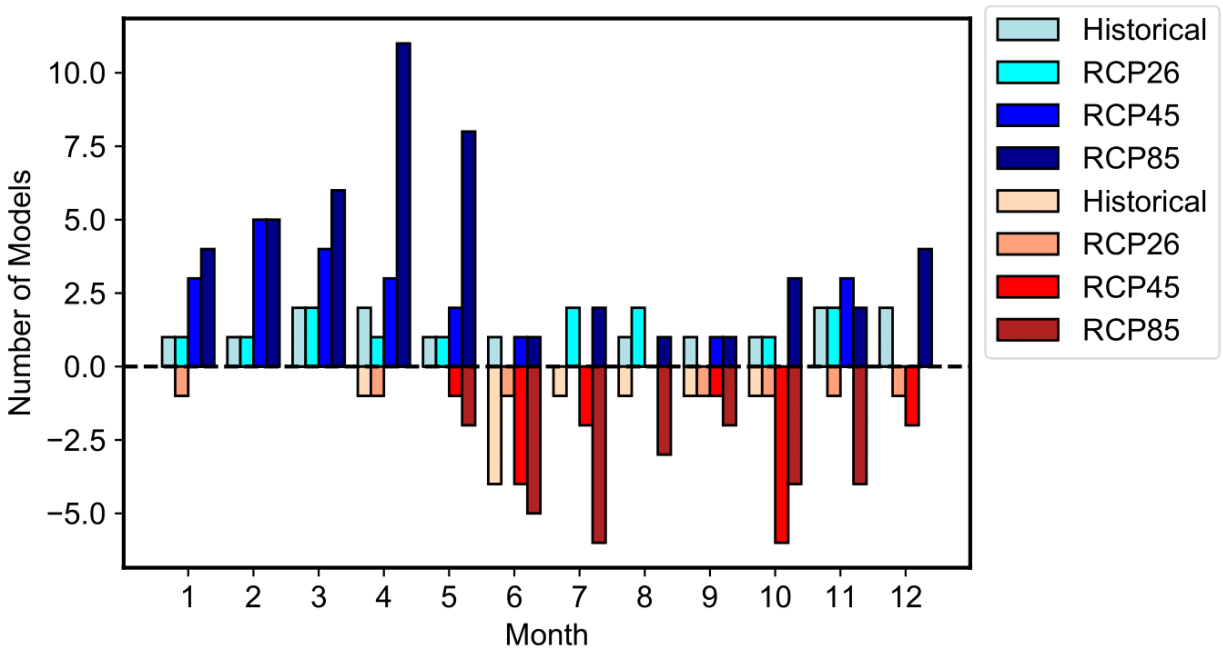

Figure 12: The number of models with Mann Kendall significance levels < 0.1 for monthly precipitation. Positive values represent models with significant positive monthly trends, and negative values represent models with significant negative trends.

The monthly precipitation trends were much less clear in the model ensemble average given in Figure 13 by the Mann Kendall 'p' values. January through April shows more increasing than decreasing trends. Averaging the Mann Kendall 'p' values appeared to drown out the models with significant trends.

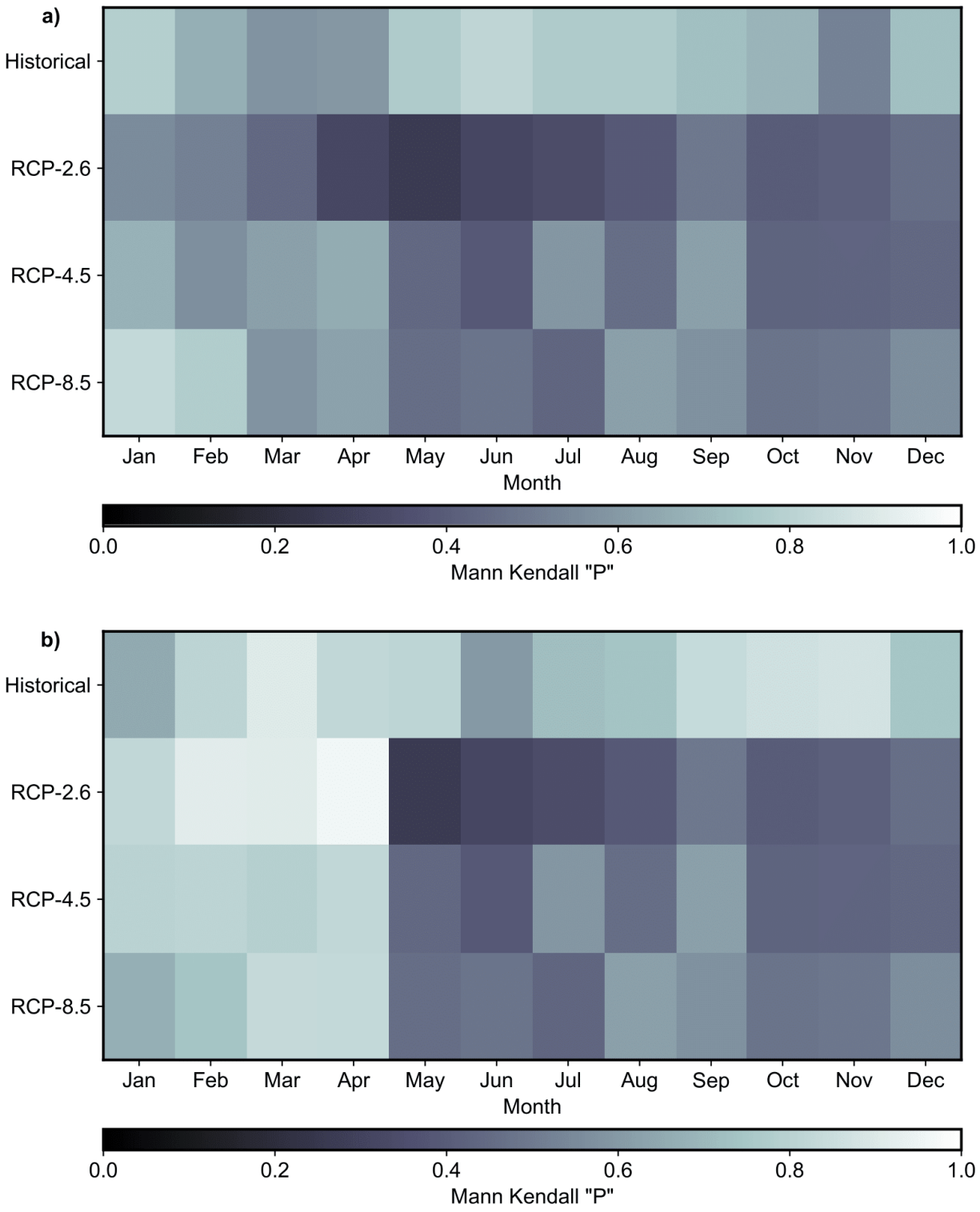

Figure 13: Mann Kendall trend test Significance levels for the model average monthly precipitation for RCP8.5, RCP4.5, and RCP2.6 for a) increasing trends and b) decreasing trends.

The gages also showed significant increasing trends for the winter (Dec.-Feb.) months and May, as well as June and November (Figure 14). July had decreasing trends across all gages. Unity Village showed a disproportionate number of increasing trends, which may be due to a significant amount of missing data that could not be infilled. Consistent changes across both the observations and models indicate increasing precipitation in the winter and spring months and decreasing precipitation during the summer.

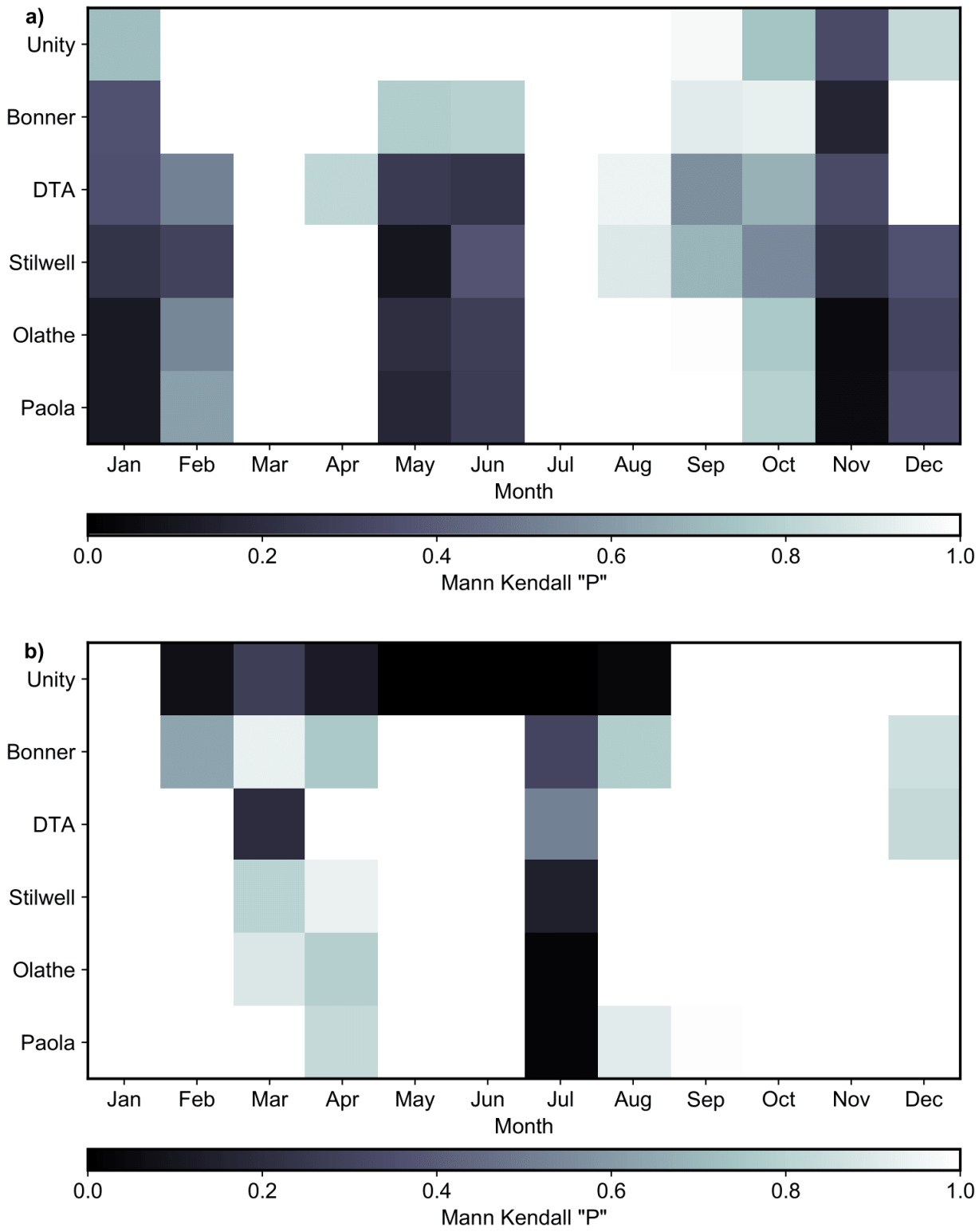

Figure 14: Mann Kendall Trend test significance levels for gage monthly precipitation, for a) increasing trends and b) decreasing trends.

3.1.2 PDS Trends

Because PDS data sets can include multiple values from a single year they are more adept at capturing the seasonal variability than AMS, which only includes the single highest event for each year. Given the seasonal changes in precipitation identified, the rest of the analysis focuses on PDS. Figure 15 shows the seasonal distribution of PDS events across all six gages for all durations. All six gages have very similar seasonal PDS patterns. The most PDS events occur in June, at about 15-17.5%, and the least fall in February at around 0-2%. The majority of PDS event occur May – October.

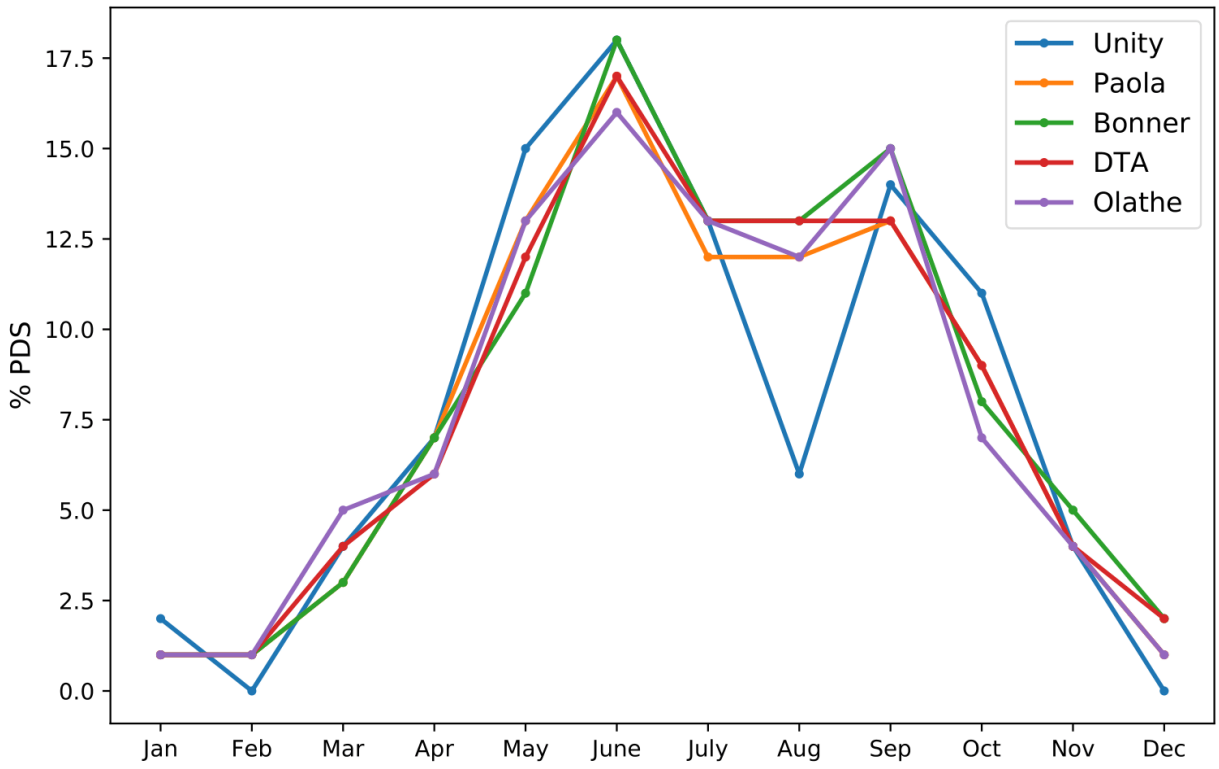

Figure 15: The seasonal distribution of PDS events for the six gages across all durations, as percentage of PDS events for each month.

The Mann Kendall trend test was used to test for significant trends in both PDS event frequency (number of events per year) and event magnitude. The results are shown in Figures 16 and 17 below for $p < 0.25$. Some negative trends were identified, but all had very low significance levels and can be considered negligible (most had $p > 0.95$). Significant increasing trends were found for both frequency and magnitude. Very few gages, however, had trends with significance levels less than 0.05.

The 24-hour event had the most gages with increasing frequency trends, while the 7-day event had the most significant frequency trend. The 12-hour event had both the most significant and the most gages with increasing magnitude trends. No gage had significant trends for both frequency and magnitude for the same storm event. Full results can be found in Figures 25 and 26 in the Appendix.

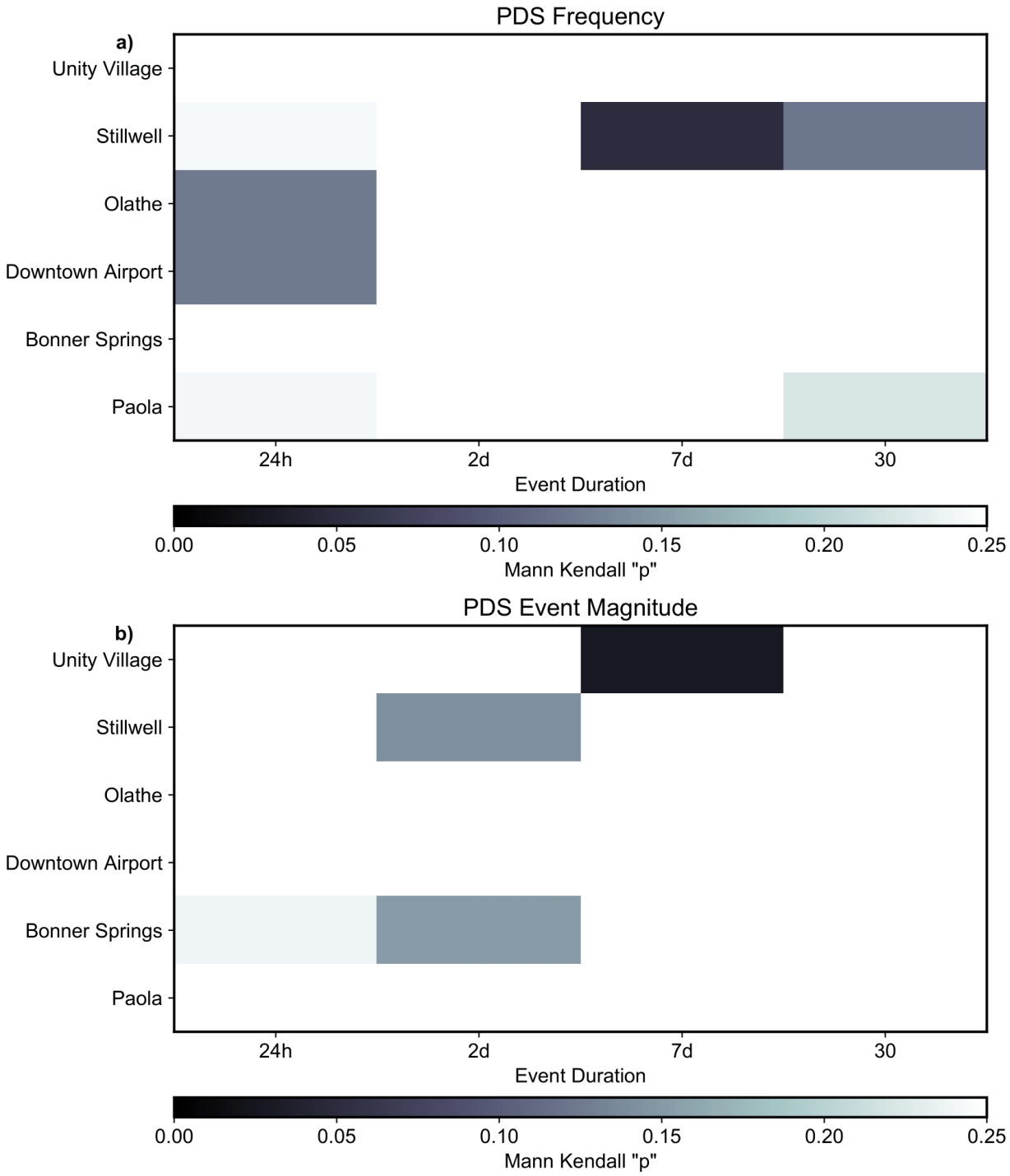

Figure 16: Mann Kendall trend test 'p' values < 0.25 for daily gage PDS where a) is the results for the number of PDS events per year and b) is the results for PDS event magnitude.

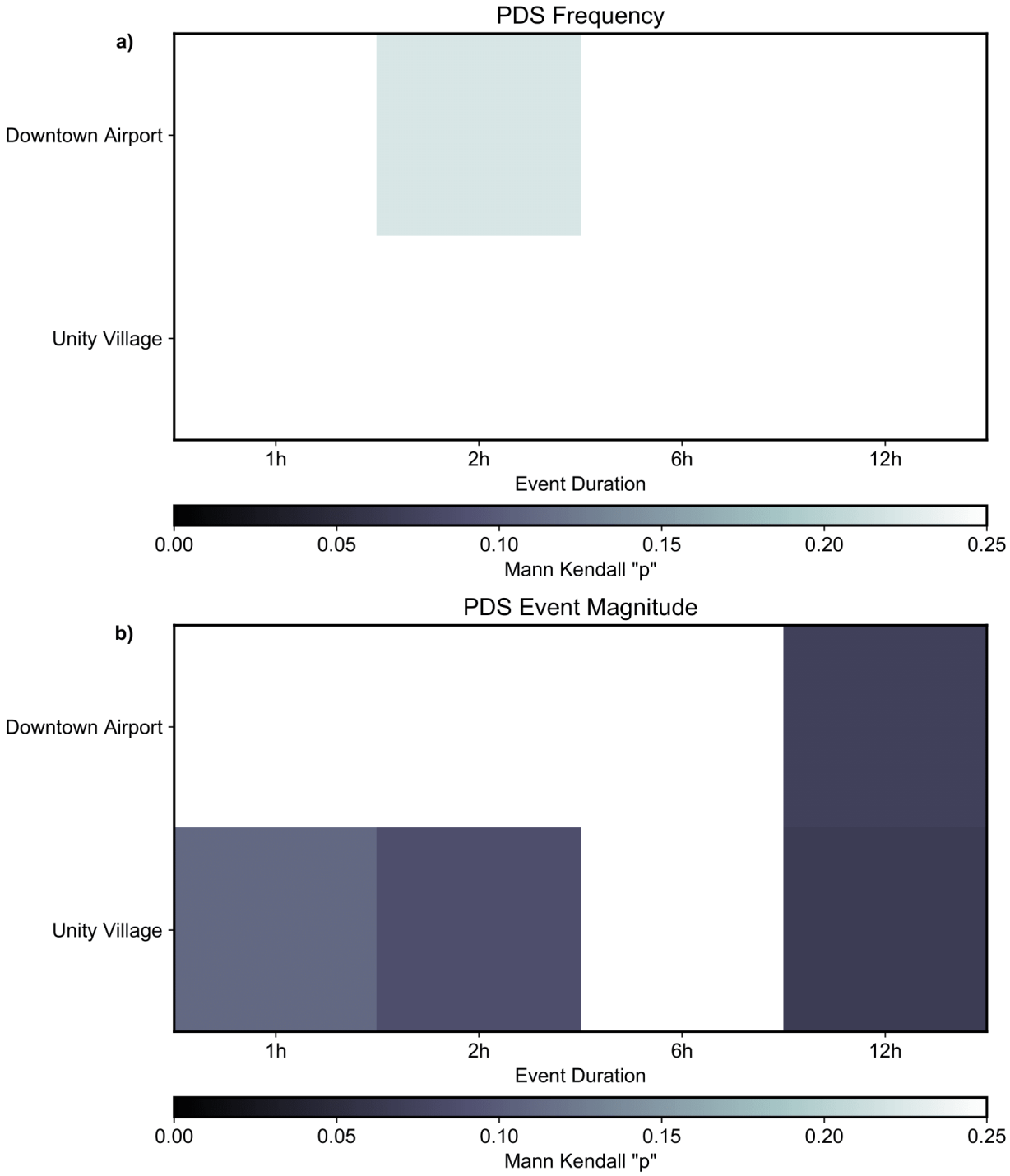

Figure 17: Mann Kendall trend test 'p' values < 0.25 for the sub daily gage PDS where a) is the results for the number of PDS events per year and b) is the results for PDS event magnitude.

3.2 Connecting Observations to Climate Model Data

The challenge with incorporating the climate model data into our IDF estimates is determining how to connect the future projections (available as monthly averages) to individual PDS storm events (at the scale of hours to days). In order to establish a relationship between the PDS events and CMIP5 projections, the relationship between the PDS data and monthly and annual precipitation is analyzed. Very little correlation was found between PDS event magnitude and annual precipitation. A stronger correlation was found at the monthly scale, but the correlation was still weak. Figure 18 shows the relationship between monthly precipitation and PDS event magnitude, with an R^2 value of only 0.37.

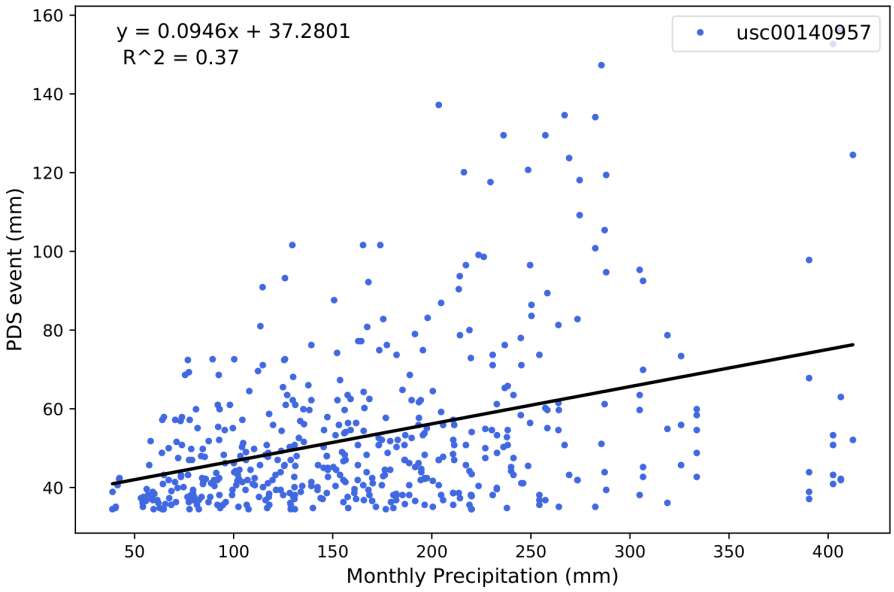

Figure 18: PDS event magnitude plotted against monthly precipitation (for the month of the event) for the Bonner Springs gage.

In contrast, a strong correlation was identified between annual exceedance rate and annual precipitation. Figure 19 shows this relationship for the 24-hour storm for all six locations. R^2 values range from 0.732 – 0.837. This correlation was identified as the best way to correlate future model changes with historical PDS data.

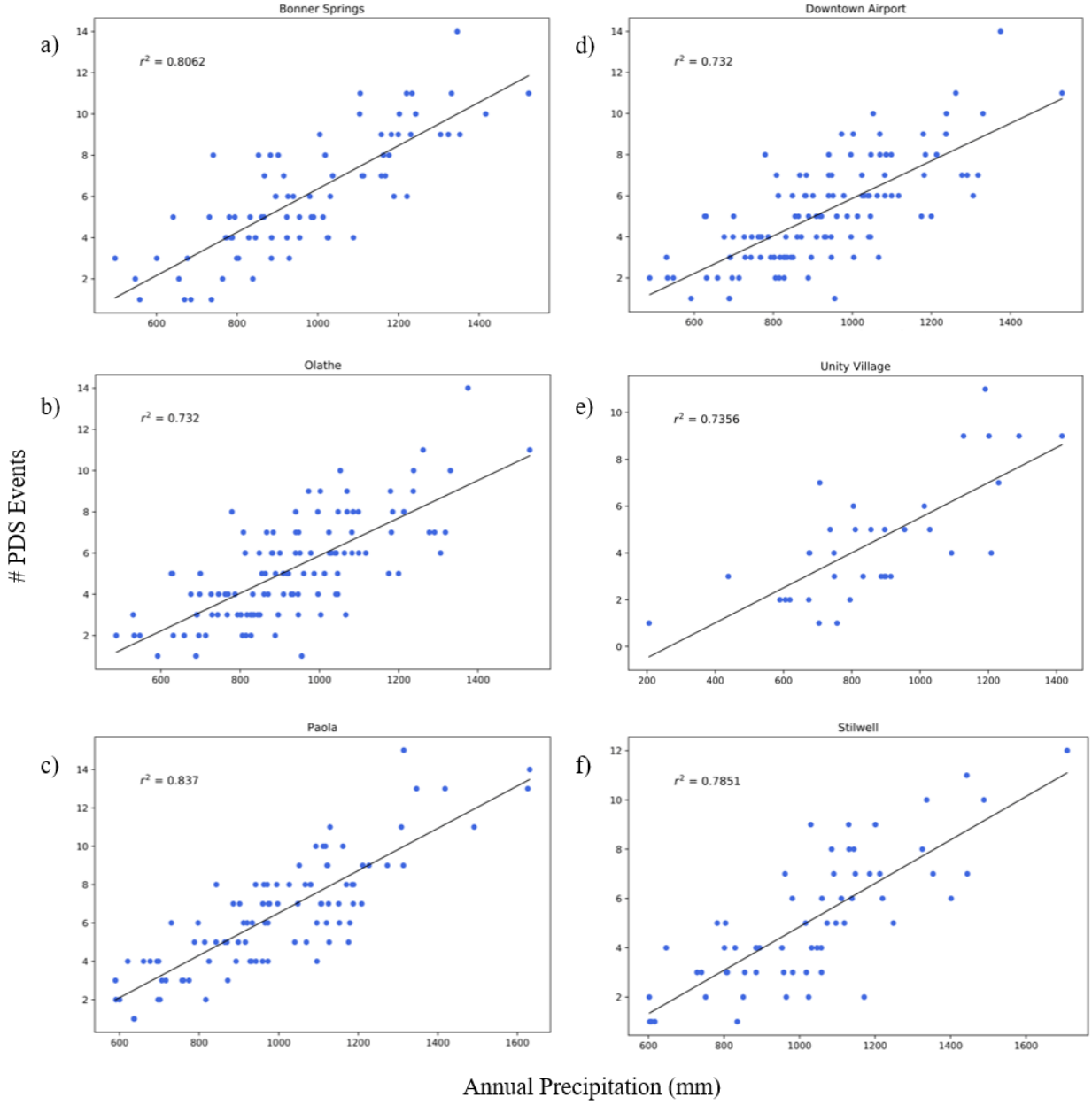

Figure 19: The relationship between annual precipitation (mm) and the number of PDS events for the a) Bonner Springs, b) Olathe, c) Paola, d) Downtown Airport, e) Unity Village, and f) Stilwell gages.

This relationship was employed to connect the future projections to the historical observations. For each individual gage and duration, a relationship between the rate of exceedance, or frequency, and

the total annual precipitation was first developed. Annual precipitation was separated into categories every 200 mm. A Frequency Occurrence Matrix was developed, containing the number of times each frequency occurred in the gage data by annual precipitation (Figure 20a). This was then converted into a Frequency Probability Matrix, showing the cumulative probability for each frequency by annual precipitation (Figure 20b).

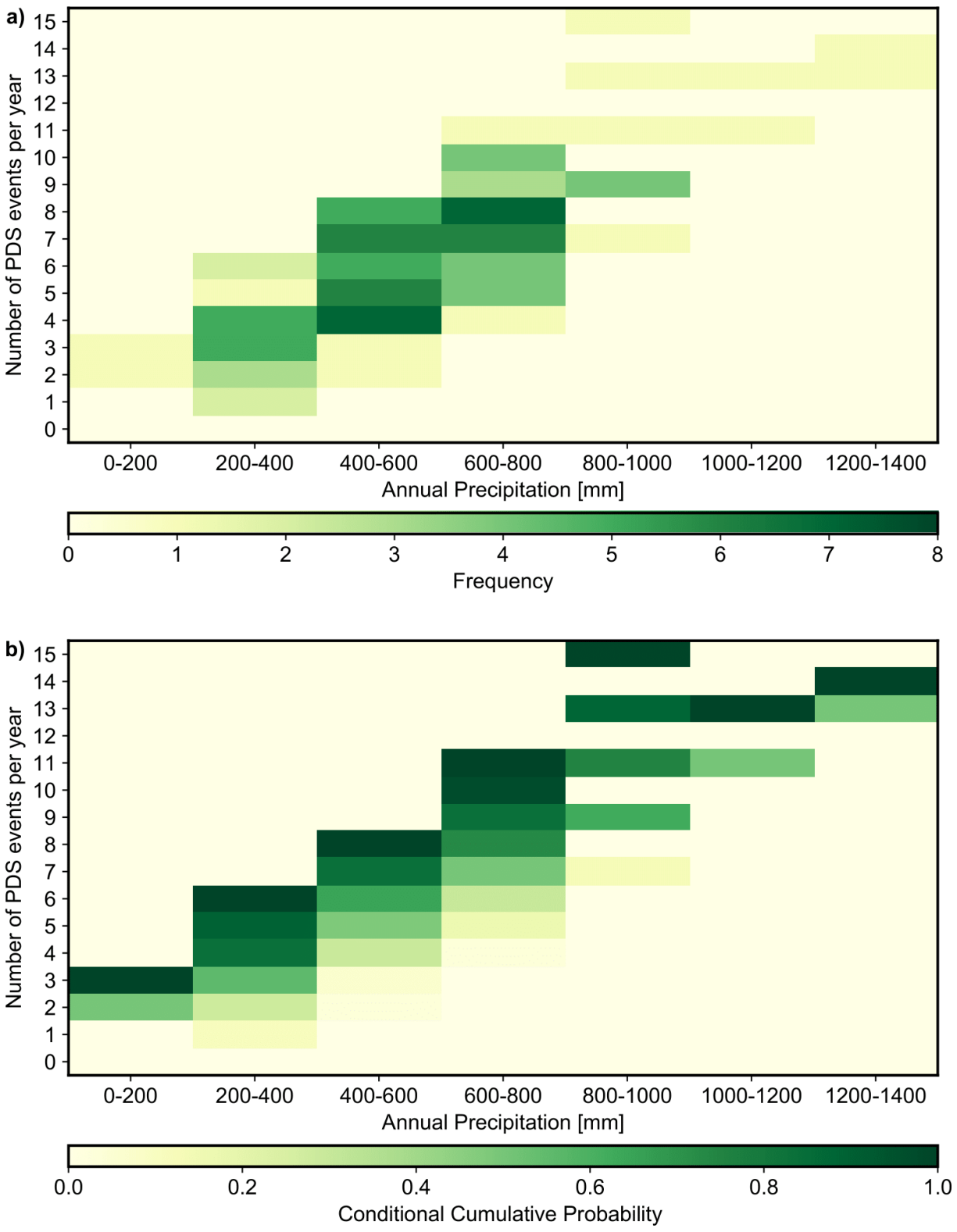

Figure 20: Example of a Frequency Occurrence Matrix, showing the number of times each frequency occurs in the observed data, grouped by annual precipitation.

This Frequency Probability Matrix was then used to predict the frequency of the projected PDS estimates. The matrix column was determined by projected annual precipitation of each climate model. Then a random number was generated between 0 and 1 and is used to choose a row within the column dictated by the climate model to determine the frequency. This was then repeated for each year (2006-2100) for each model.

Once the frequency of PDS values for each future year is determined, a distribution of PDS values fit for each individual gage and duration was used to select PDS values for each future year by randomly selecting from the distribution. This then produces a projected PDS covering 2005-2100. This synthetic projected distribution of PDS values is then fit to either a GEV or GP. The new distribution is then used to estimate storm intensity for the 2, 5, 10, 25, 50, and 100-year events. The synthetic nature of this methodology requires an ensemble approach. This is done by running the above process 1000 times for each CMIP5 model ensemble member to produce X different storm intensity estimates for each duration (1-hour, 5-hour, …). This large ensemble was then used to determine the final IDF curves using the average across all ensemble members as well as the confidence intervals using the 'bootstrapped' library in python. A flow chart of this process is shown in Figure 21 below.

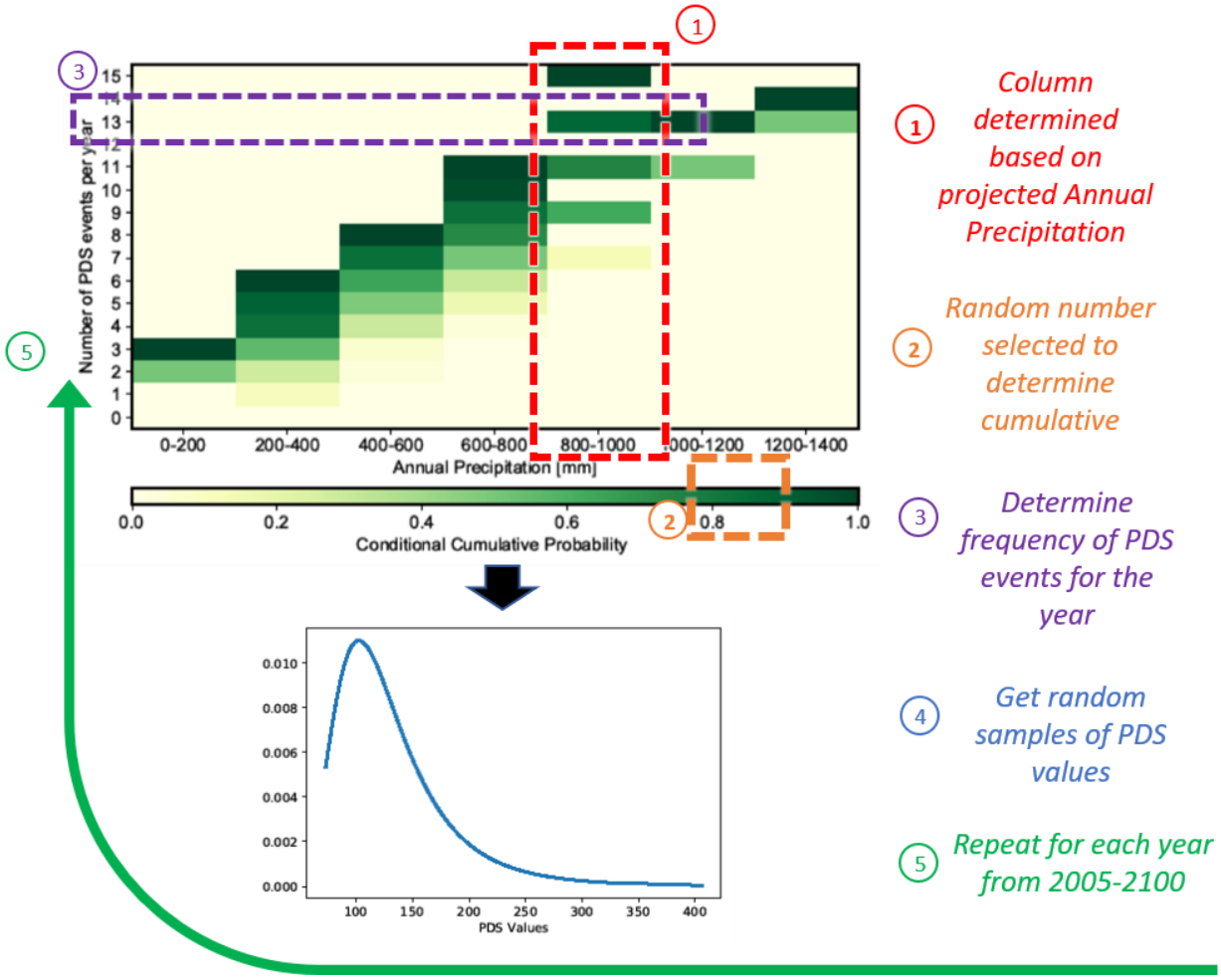

Figure 21: Flow chart describing the process of developing new PDS data sets using the frequency probability matrix.

The method above was used to develop adjusted PDS data sets were through the year 2100 incorporating the climate model information. The relationship between PDS event frequency and annual precipitation was used to determine future event frequency. The historical PDS distribution was then used to determine event magnitudes. This new PDS was then used to develop IDF estimates, using NOAA's methods.

3.3 IDF Estimates for Future Change

Using the previously described method, future changes in precipitation are used to develop projected IDF estimates. The results show an increase in precipitation IDF estimates from the original for most durations and recurrence intervals, across all gages. Figure 22 compares the projected estimates to the estimates based on gage data only. The largest change can be seen in the 100-year event, while the 5, 10, and 20-year events have much smaller changes. Confidence intervals were developed using bootstrapping for uncertainty between model ensemble members (Figures 28-32 in the appendix). Overall, there was very strong agreement between the model ensemble members. The 100-year event tended to have the largest confidence interval, while the 2-year event had the smallest.

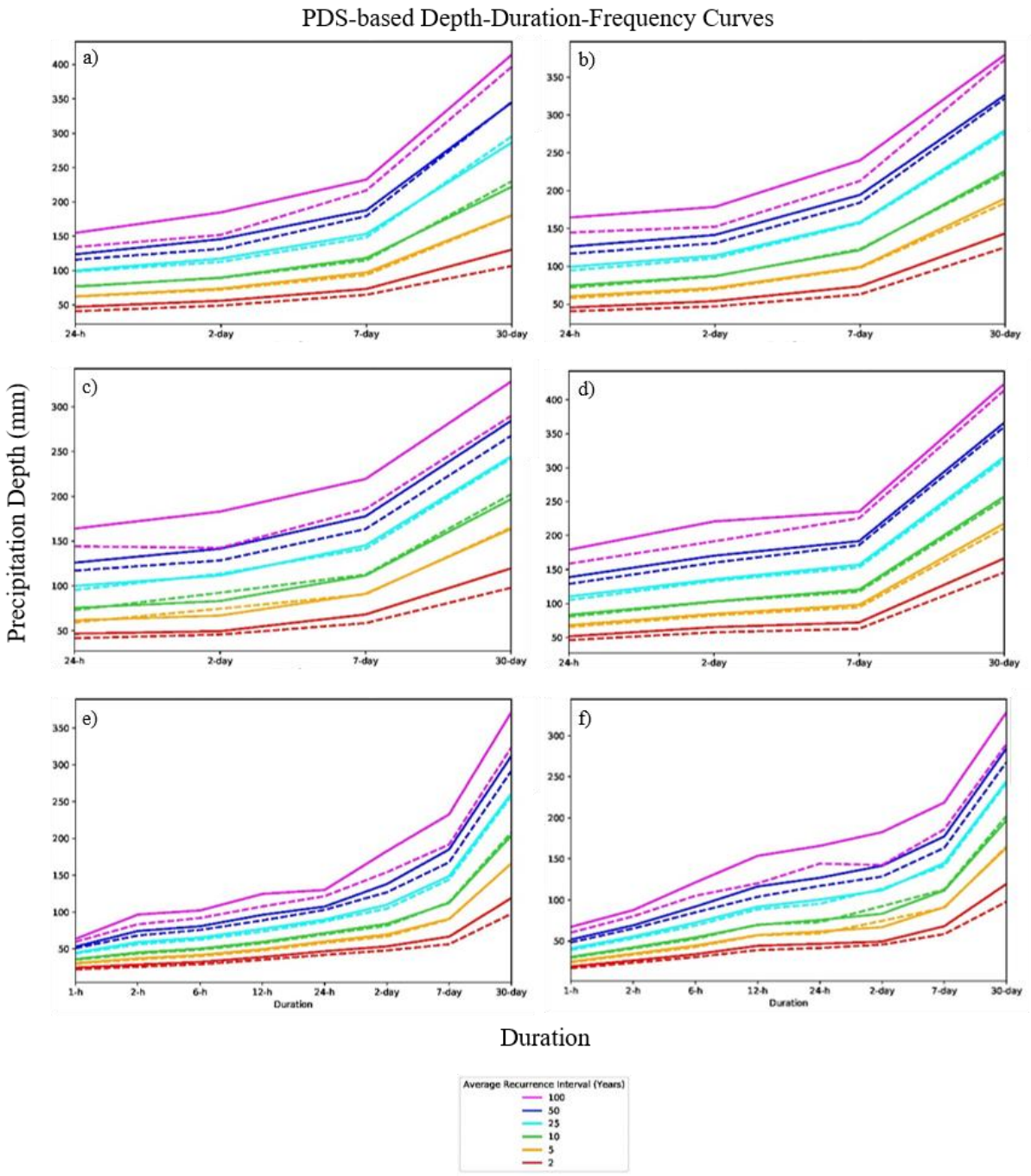

Figure 22: Depth-Duration Frequency Curve estimates for Bonner Springs (a), Paola (b), Olathe (c), Stilwell (d), Unity Village (e), and Downtown Airport (f). The dashed lines are estimates based only on historical data. The solid lines are estimates incorporating CMIP5 projections.

The percent change was calculated for each event (Figures 23 and 24). There was good agreement between the gages. The two-year and 100-year event increased by the largest margin for all gages and durations. It is expected that the 2-year would increase substantially due to the higher frequency of intense events. When averaged across all gages, the 100-year events increased by 11-20%. The 2-year events increased by 7-22%.

The 5, 20, and 25-year events had much smaller changes. The gage average ranged from -2-10%. The percentage change for the 50-year event was generally in the middle. The gage average was -2-7%. Although there were some negative changes, most were small enough to be considered insignificant. The most significant negative change was the 2-day storm for the Olathe and Downtown Airport gages, with decreases of 10%. Full results can be found in Tables 17-22 in the Appendix.

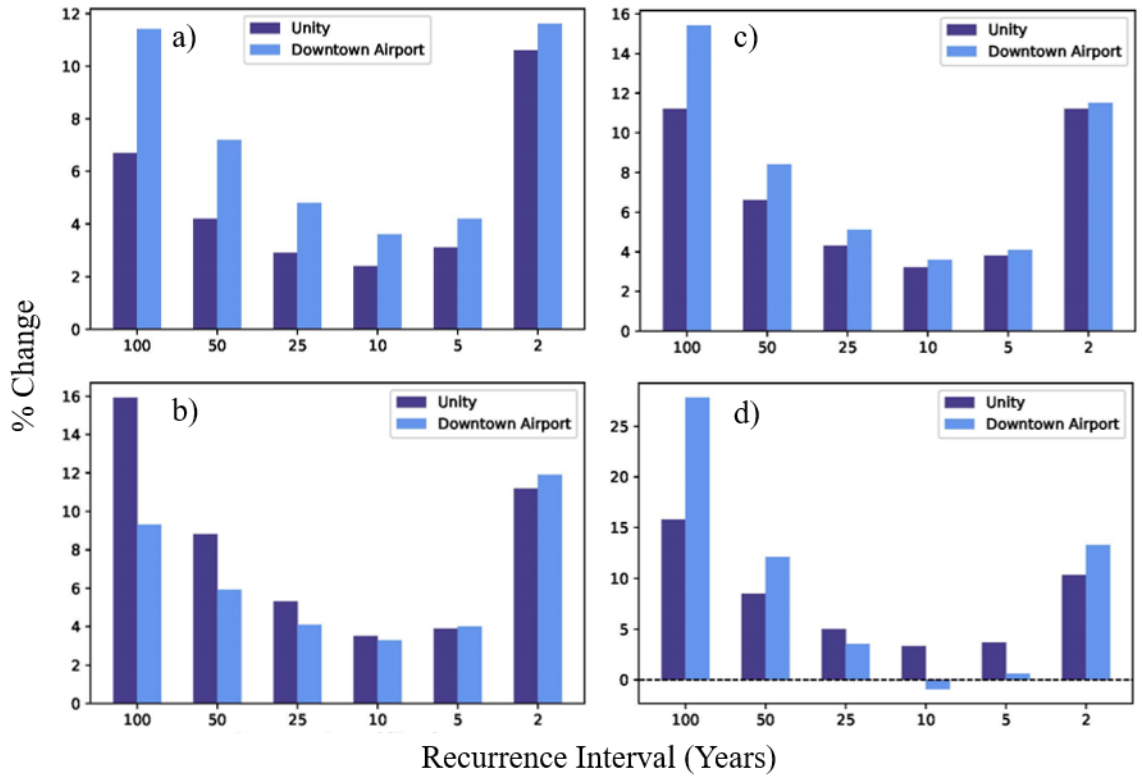

Figure 23: The percent change in event magnitude for the 1-hour (a), 2-hour (b), 6-hour (c), and 12-hour events.

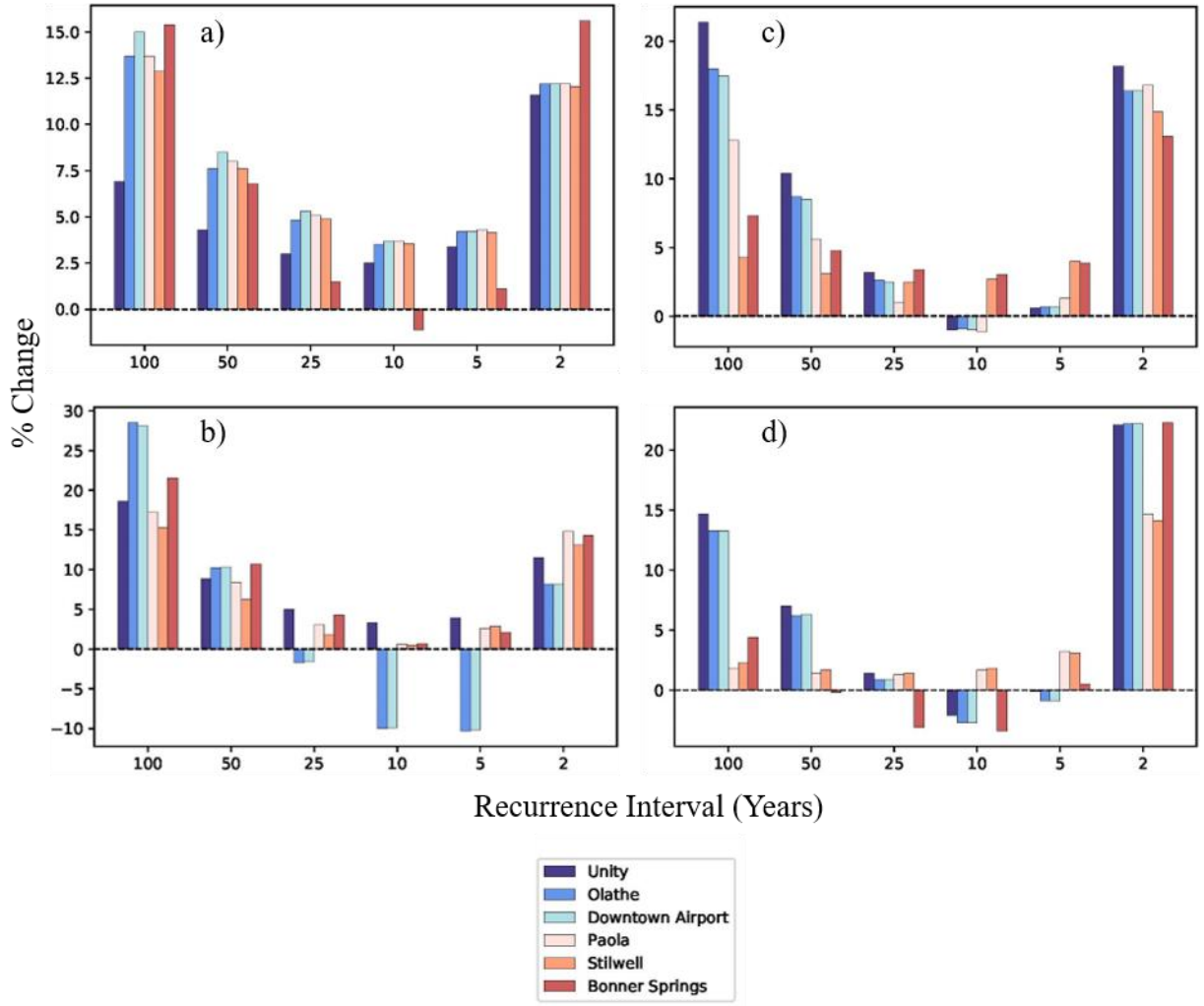

Figure 24: The percent change in event magnitude for the 24-hour (a), 2-day (b), 7-day (c), and 30-day (d) events.

Chapter 4: Summary and Conclusions

The purpose of this study was to characterize projected changes in storm intensity, duration, and frequency for the Kansas City area using historical gage data and the CMIP5 model ensemble. Precipitation trends were analyzed for annual and seasonal precipitation for gage and model data. Trends in PDS event precipitation were analyzed for gage data. IDF curves were then developed incorporating the climate model projections.

Model bias was found, which varied widely across the 20 models and 6 gages. This model bias was corrected using CDF matching for each gage, to provide better representation from the climate models. Precipitation trends were then analyzed for the models and gages. Analyzing both annual and seasonal trends for both the model and gage data allowed the use of future model information (available as monthly averages) to glean information about potential changes in extreme precipitation events.

Annual precipitation could be compared directly between the climate models and gage data. Significant increasing trends were identified in annual precipitation for both the gage data and CMIP5 model data When averaged across models, both the 'p' values were similar for both positive and negative trends. However, there were more models across all RCPs with increasing trends (59 models) than decreasing trends (23 models). More significance was found for increasing trends than decreasing trends as well. RCP 8.5 had the most increasing trends. Increasing trends in annual precipitation were also identified for all 6 gages, with p values ranging from 0.02-0.077.

Seasonal precipitation trends were then analyzed, to further identify alterations in expected precipitation patterns. For both the models and gages, the strongest increasing signals were found in the winter (Dec-Feb) and spring (Mar-April) months, with some decreasing trends from June to August or September. The Mann Kendall results showed a significance level of $p < 0.1$ for only some of the gages and models, but there was strong agreement across all gages and models for the directions of the seasonal trends. Historical gage data was then analyzed for significant trends in the PDS data sets using the Mann

Kendall trend test. While significant trends were not identified for some PDS events, increasing trends in both the magnitude and frequency of many PDS events were found.

A strong correlation was identified between PDS event frequency and annual precipitation. This relationship was used to relate information from climate model projections at the monthly scale to gage-based PDS events. The yearly PDS frequency for the future time period was determined based on the CMIP5 model projected annual precipitation. IDF estimates incorporating projected climate model trends were then developed using the adjusted PDS data sets.

This information on trends and the relationship between PDS event frequency and annual precipitation were employed to develop IDF estimates for all 6 gages for the 24-hour, 2-day, 7-day, and 30-day storms, and the 1-hour, 2-hour, 6-hour, and 12-hour storms for the hourly gages, with recurrence intervals of 2, 5, 10, 25, 50, and 100 years. Results showed an increase in event magnitude from the original for most durations and recurrence intervals, across all gages. The 2-year and 100-year events increased the most for all gages and durations, while the 5-year and 10-year events changed only marginally (and in some cases marginally decreased).

More significant increases were found for the 2-year and 100-year events than the other frequency storms. It is possible that this is an artifact of the process used for estimation rather than a true reflection of trends. The sample of randomly generated numbers may be skewed toward zero and one, leading to an increased frequency in both the lower (100-year) and higher (2-year) events. This could be tested by increasing the number of runs and randomly generated numbers. The trend could also be caused by a skew in the distributions. Since the GEV and GP distributions are heavy-tailed, an increase in the frequency of the lower probability events could lead to a greater increase in the resulting event magnitude, compared to the higher probability events that do not fall into the upper tail range. More analysis is needed to determine if the greater increase in the 2 and 100-year events is a true reflection of trends from the data.

Based on the results, the Kansas City area is likely to see more annual rainfall going forward. However, that does not necessarily indicate having more available water. Significant increases in the 2-year and 100-year events would indicate that both the more frequent extreme events and the rarer most intense events will increase in intensity, dropping more of the rainfall in a shorter period of time. These intense storm events lead to rapid runoff and a smaller percentage of rainfall captured within the system. This can have broad implications, including increased flooding, a decreased ability to store water in reservoirs, and increased hydromodification and sedimentation.

Seasonality also plays an important role in the effects of precipitation. The results showed increasing trends in the winter (Dec-Feb) and spring (Mar-Apr) months, while the majority of PDS events occur between May and September. This could indicate a potential shift in seasonal precipitation patterns for the Kansas City area, including the potential of drier summer months. Although the Kansas City area is mostly urbanized, surrounding rural farming areas could also be strongly impacted by the altered seasonality of precipitation.

An increase in the magnitude of the 100-year event has serious implications for engineering design. Critical infrastructure, such as bridges, roads, and overflow channels, that are designed for the 100-year event using stationary methods may be under designed. Increased smaller events, such as the 2-year storm, can also affect engineering designs. Peak runoff control is designed for multiple storms, including the 2-year event (APWA, 2011). Erosion protection measures are also designed for the 2-year event (APWA, 2011). Systems like this, which are critical to overall watershed health, may also be under designed. Engineers and planners should consider looking beyond the historical IDF estimates to develop more robust and sustainable designs.

This study only analyzed precipitation data and because a lack of correspondence between precipitation and streamflow often exists (Sharma, Wasko, & Lettenmaier, 2018), changing precipitation patterns could have a wide array of effects on streamflows. The streamflows not only depend on seasonal timing, but also on the current environmental conditions, such as antecedent moisture condition and

moisture convergence. Urban catchments, however, are generally precipitation driven (Sharma, Wasko, & Lettenmaier, 2018), so flood peaks will likely increase with the increased precipitation events within the urbanized areas of Kansas City. Further analysis is needed to determine streamflow scenarios that could result from changing precipitation patterns.

In the future, regional frequency analysis can be used to develop more robust historical data sets for the Kansas City area with the ability to use data from more gages with limited periods of record. Distributions with nonstationary parameters could also be developed for the area and compared with the results of this study. Additionally, potential changes in streamflow patterns can be analyzed using the precipitation data.

Chapter 5: Appendix

Table 11: Regression Information, the percentage of missing data filled in by each individual regression.

	Missing Original	% Filled In				Missing Remaining
		Regression #1	Regression #2	Regression #3	Regression #4	
Unity	38%	22%	3%	0.2%	-	0%
Stillwell	8%	7%	2%	-	-	0%
Olathe	9%	1%	1%	4%	2%	0.3%
Bonner Springs	0.6%	0.6%	0.03%	-	-	0.003%
Paola	14%	0.5%	11%	2%	-	0.1%
DTA	6%	5%	1%	0.01%	-	0%

Table 12: Gages used to fill in missing data using linear regression.

Gage y	**Gage X**			
	Regression #1	**Regression #2**	**Regression #3**	**Regression #4**
Unity	Stanley	Swope Park	Downtown Airport	
Stilwell	Olathe	Paola		
Olathe	Stilwell	Bonner Springs	Paola	Downtown Airport
Bonner	Olathe	Independence		
Paola	Stilwell	Olathe	Downtown Airport	
DTA	Gladstone	International Airport	Olathe	

Table 13: Winter R^2 values for each linear regression.

Gage	R^2			
	Regression #1	**Regression #2**	**Regression #3**	**Regression #4**
Unity	0.27	0.41	0.22	-
Stilwell	0.66	0.63	-	-
Olathe	0.66	0.63	0.62	0.24
Bonner	0.63	0.73	-	-
Paola	0.63	0.62	0.21	-
Downtown Airport	0.79	0.48	0.24	-

Table 14: Fall R^2 values for each linear regression.

Gage	R^2			
	Regression #1	**Regression #2**	**Regression #3**	**Regression #4**
Unity	0.2	0.2	0.23	-
Stilwell	0.7	0.64	-	-
Olathe	0.7	0.59	0.49	0.24
Bonner	0.59	0.59	-	-
Paola	0.64	0.49	0.17	-
Downtown Airport	0.79	0.58	0.24	-

Table 15: Spring R^2 values for each linear regression.

Gage	R^2			
	Regression #1	**Regression #2**	**Regression #3**	**Regression #4**
Unity	0.35	0.28	0.14	-
Stilwell	0.64	0.53	-	-
Olathe	0.64	0.49	0.45	0.21
Bonner	0.49	0.51	-	-
Paola	0.53	0.45	0.13	-
Downtown Airport	0.83	0.54	0.21	-

Table 16: Summer R^2 values for each linear regression.

Gage	R^2			
	Regression #1	**Regression #2**	**Regression #3**	**Regression #4**
Unity	0.11	0.2	0.11	-
Stilwell	0.5	0.32	-	-
Olathe	0.5	0.36	0.24	0.14
Bonner	0.36	0.33	-	-
Paola	0.32	0.24	0.03	-
Downtown Airport	0.78	0.38	0.24	-

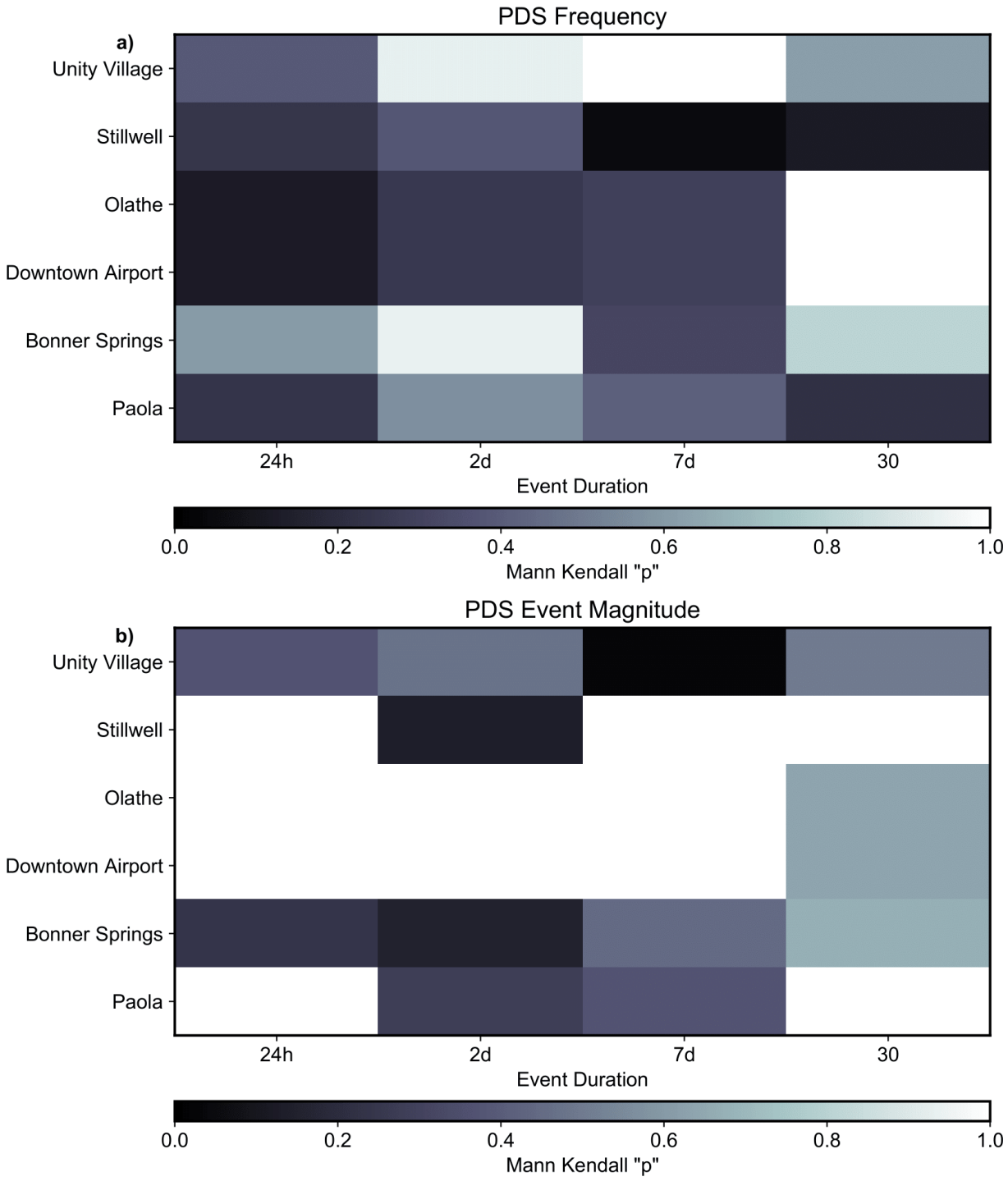

Figure 25: Mann Kendall trend test 'p' values for the gages daily PDS where a) is the results for the number of PDS events per year and b) is the results for PDS event magnitude.

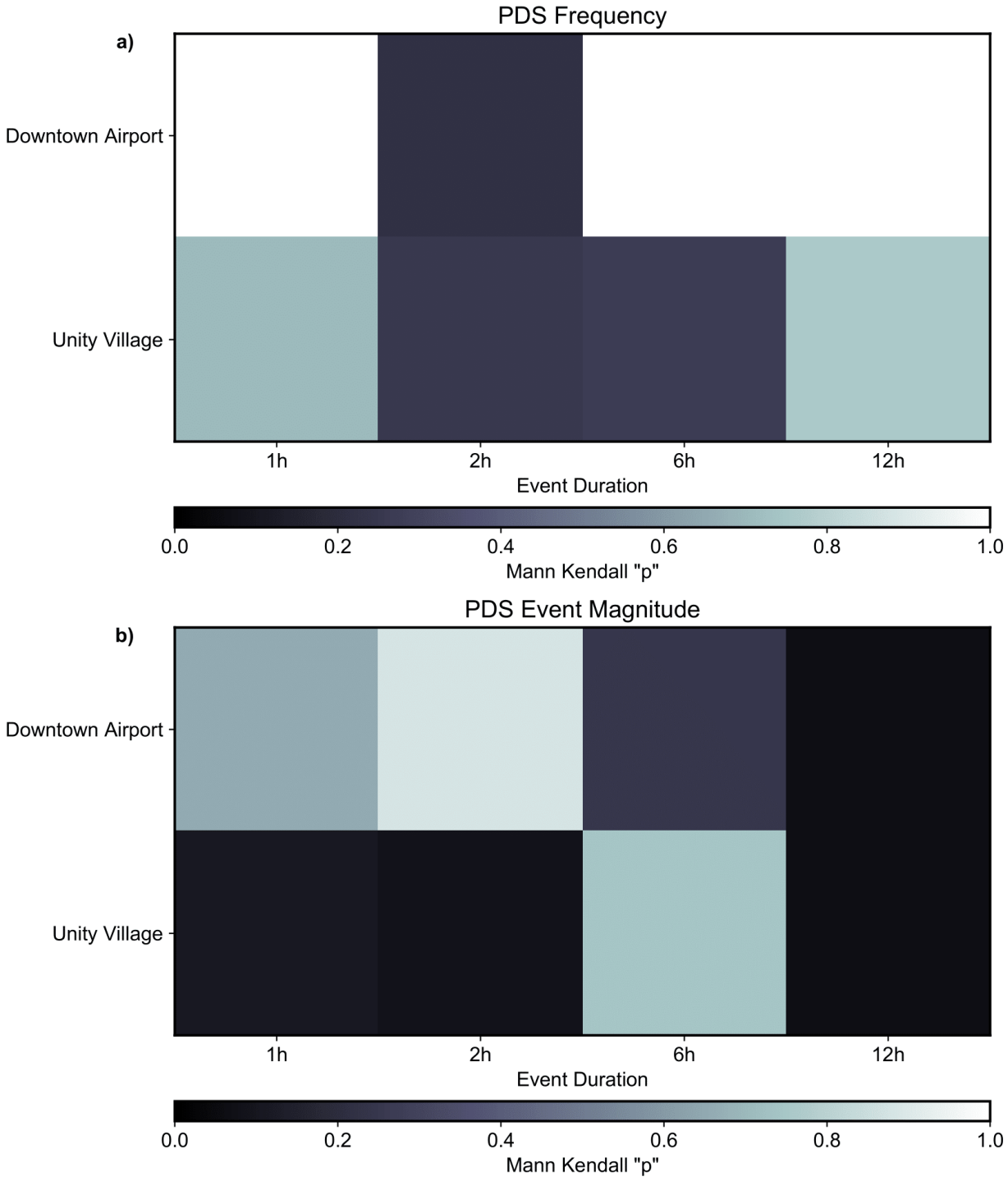

Figure 26: Mann Kendall trend test 'p' values for the gages sub daily PDS where a) is the results for the number of PDS events per year and b) is the results for PDS event magnitude.

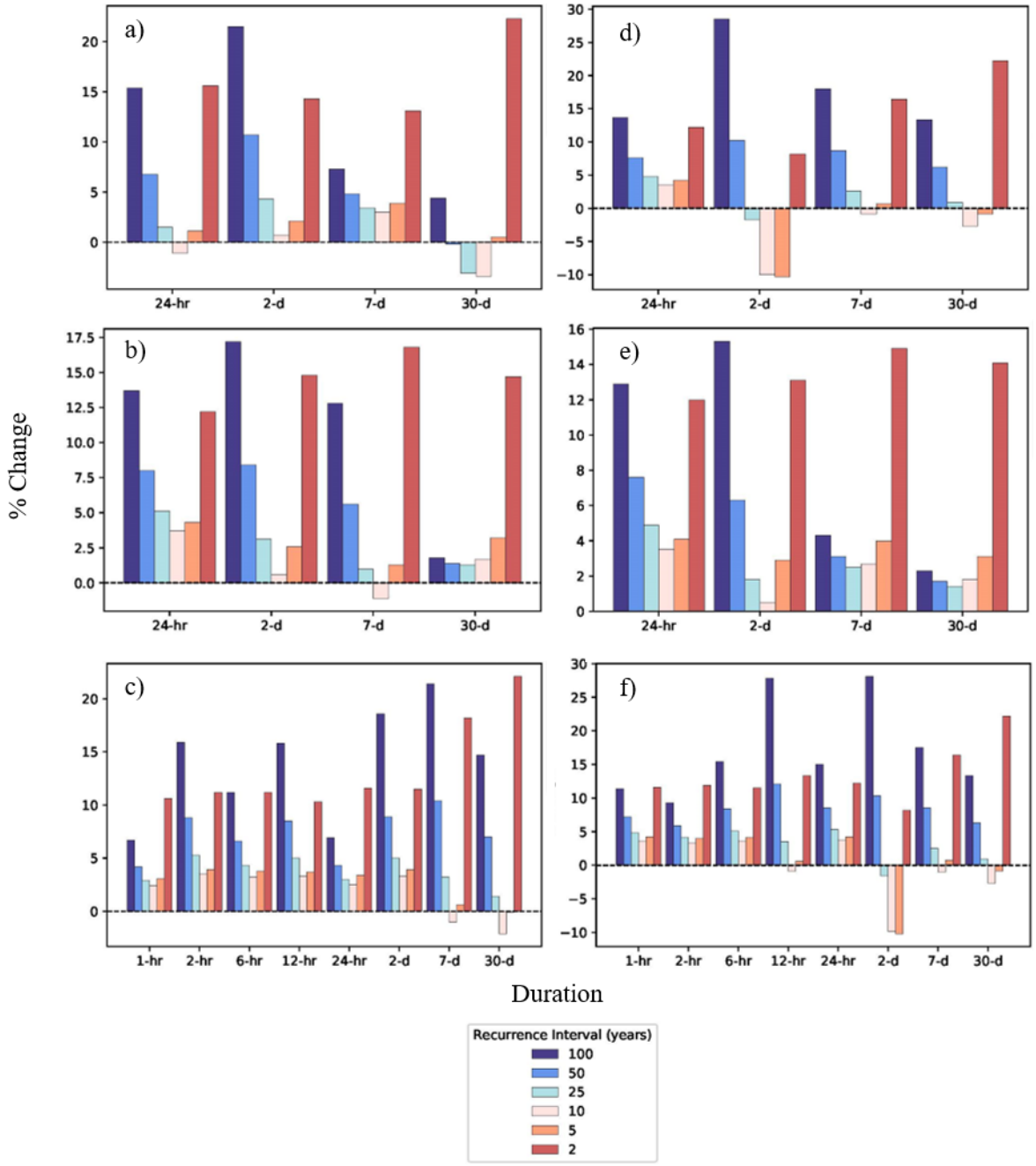

Figure 27: The percent change in event magnitude for the a) Bonner Springs, b) Paola, c) Unity Village, d) Olathe, e) Stilwell, and f) Downtown Airport gages.

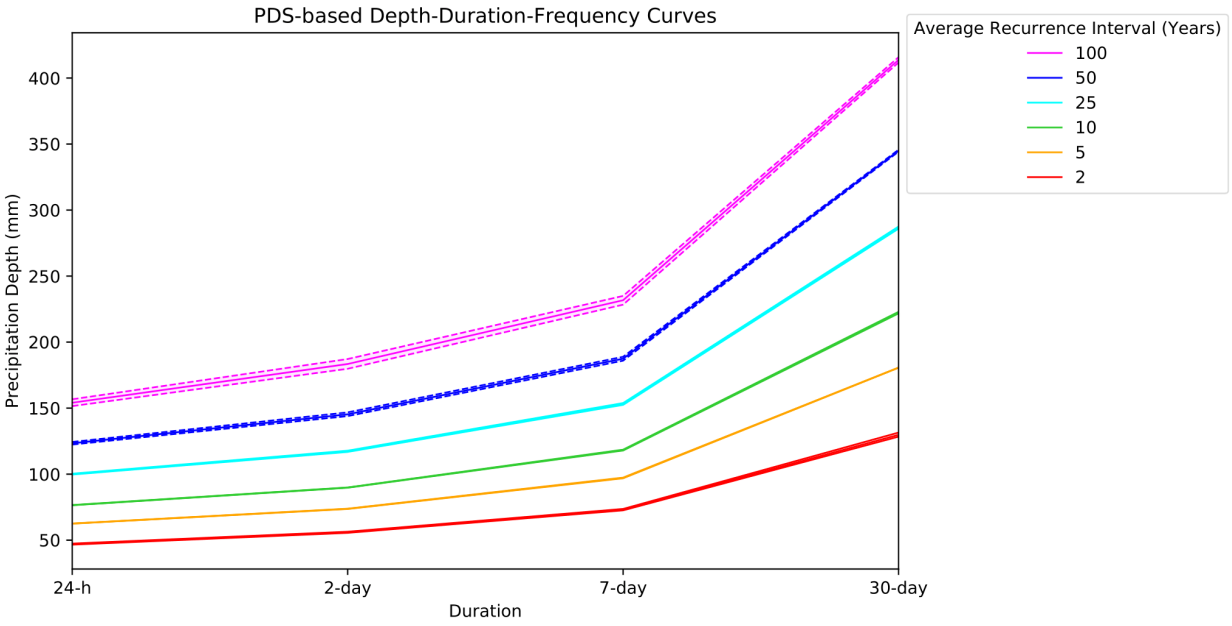

Figure 28: IDF curves for Bonner Springs with confidence intervals.

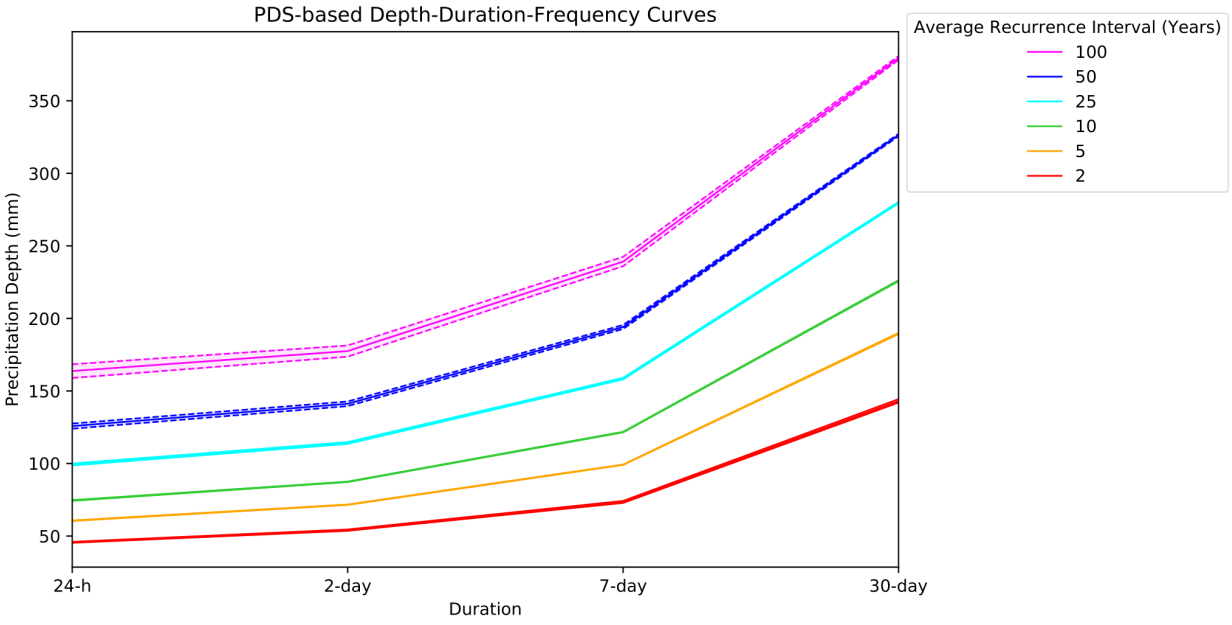

Figure 29: IDF curves for Paola with confidence intervals.

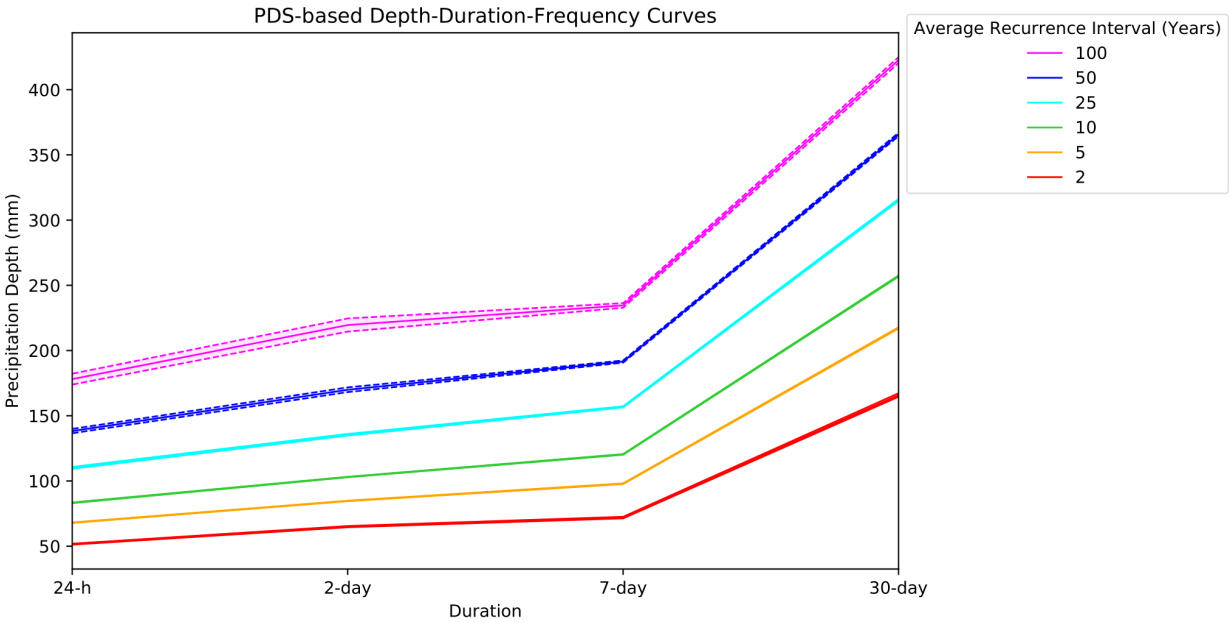

Figure 30: IDF curves for Stilwell with confidence intervals.

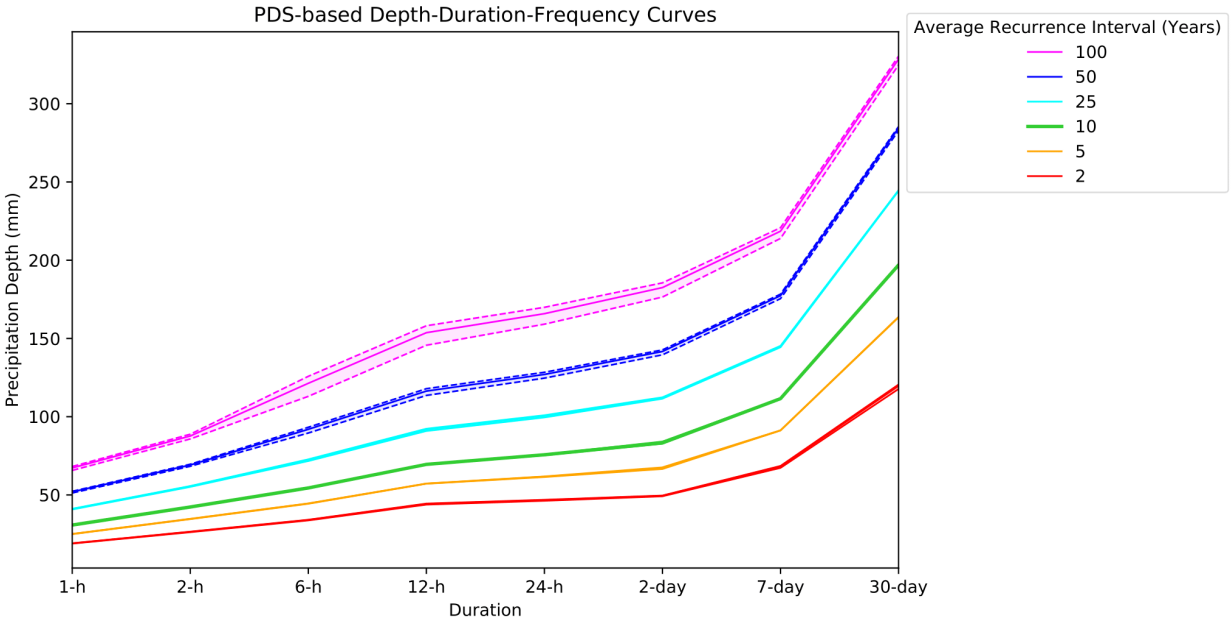

Figure 31: IDF curves for Downtown Airport with confidence intervals

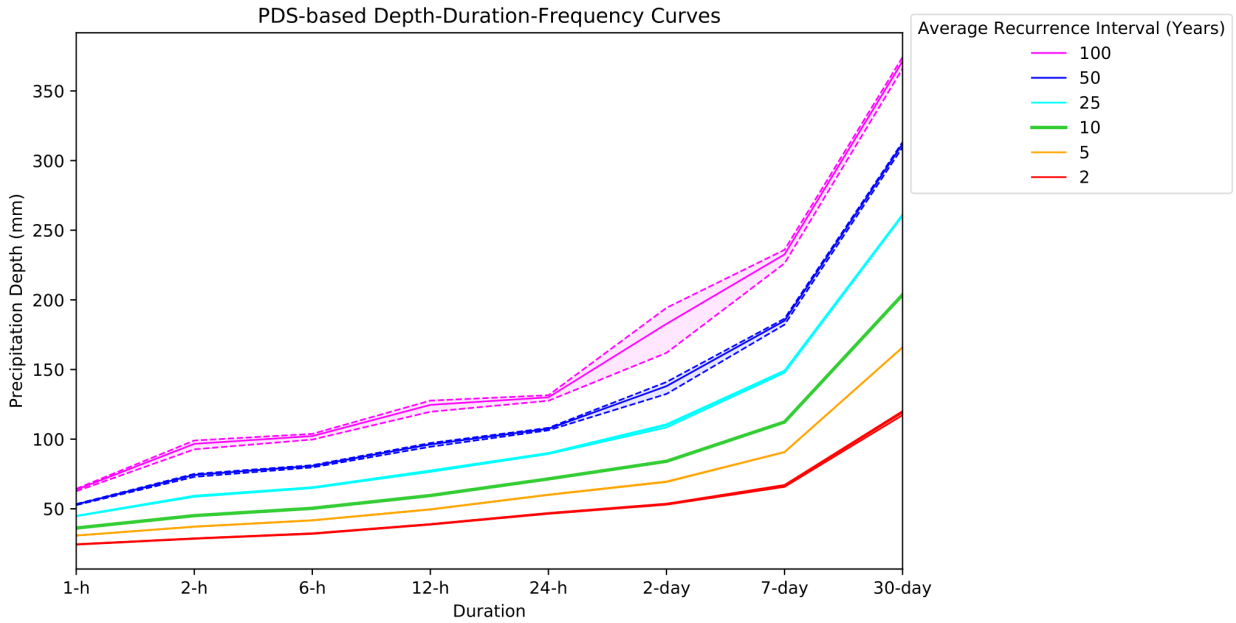

Figure 32: IDF curves for Unity Village with confidence intervals.

Table 17: The percent change in IDF event magnitude, averaged across all gages.

Recurrence Interval	Duration							
	1-hr	2-hr	6-hr	12-hr	24-hr	2-day	7-day	30-day
2	7%	13%	13%	22%	13%	22%	14%	8%
5	5%	7%	7%	10%	7%	9%	7%	4%
10	3%	5%	5%	4%	4%	2%	3%	0
25	3%	3%	3%	1%	3%	-2%	0%	-1
50	7%	4%	4%	2%	4%	-2%	2%	1
100	11%	12%	11%	12%	13%	12%	16%	20

Table 18: The percent change in IDF event magnitude for the Stilwell gage.

Duration	Frequency					
	100	50	25	10	5	2
24 hour	13%	8%	5%	4%	4%	12%
2 day	15%	6%	2%	1%	3%	13%
7 day	4%	3%	3%	3%	4%	15%
30 day	2%	2%	1%	2%	3%	14%

Table 19: The percent change in IDF event magnitude for the Paola gage.

Duration	Frequency					
	100	50	25	10	5	2
24 hour	14%	8%	5%	4%	4%	12%
2 day	17%	8%	3%	1%	3%	15%
7 day	13%	6%	1%	-1%	1%	17%
30 day	2%	1%	1%	2%	3%	15%

Table 20: The percent change in IDF event magnitude for the Olathe gage.

Duration	Recurrence Interval (years)					
	100	50	25	10	5	2
24 hour	14%	8%	5%	4%	4%	12%
2 day	28%	10%	-2%	-10%	-10%	8%
7 day	18%	9%	3%	-1%	1%	16%
30 day	13%	6%	1%	-3%	-1%	22%

Table 21: The percent change in IDF event magnitude for the Downtown Airport gage.

Duration	Recurrence Interval (years)					
	100	50	25	10	5	2
1 hour	11%	7%	5%	4%	4%	12%
2 hour	9%	6%	4%	3%	4%	12%
6 hour	15%	8%	5%	4%	4%	11%
12 hour	28%	12%	4%	-1%	1%	13%
24 hour	15%	8%	5%	4%	4%	12%
2 day	28%	10%	-2%	-10%	-10%	8%
7 day	18%	8%	2%	-1%	1%	16%
30 day	13%	6%	1%	-3%	-1%	22%

Table 22: The percent change in IDF event magnitude for the Unity Village gage.

Duration	Recurrence Interval (years)					
	100	50	25	10	5	2
1 hour	7%	4%	3%	2%	3%	11%
2 hour	16%	9%	5%	4%	4%	11%
6 hour	11%	7%	4%	3%	4%	11%
12 hour	16%	9%	5%	3%	4%	10%
24 hour	7%	4%	3%	3%	3%	12%
2 day	19%	9%	5%	3%	4%	12%
7 day	21%	10%	3%	-1%	1%	18%
30 day	15%	7%	1%	-2%	0%	22%

Chapter 6: References

Allan, R. P., & Soden, B. J. (2008). Atmospheric warming and the amplification of precipitation extremes. *Science*, *321*(5895), 1481–1484. https://doi.org/10.1126/science.1160787

Bureau of Reclamation. (2013). *Downscaled CMIP3 and CMIP5 Climate and Hydrology Projections*. Retrieved from https://gdo-dcp.ucllnl.org/downscaled_cmip_projections/

Buytaert, W., Célleri, R., & Timbe, L. (2009). Predicting climate change impacts on water resources in the tropical Andes: Effects of GCM uncertainty. *Geophysical Research Letters*, *36*(7), 1–5. https://doi.org/10.1029/2008GL037048

Chandra, R., Saha, U., & Mujumdar, P. P. (2015). Model and parameter uncertainty in IDF relationships under climate change. *Advances in Water Resources*, *79*, 127–139. https://doi.org/10.1016/j.advwatres.2015.02.011

Condon, L. E., Gangopadhyay, S., & Pruitt, T. (2015). Climate change and non-stationary flood risk for the upper Truckee River basin. *Hydrology and Earth System Sciences*, *19*(1), 159–175. https://doi.org/10.5194/hess-19-159-2015

Dessai, S., & Hulme, M. (2013). Does climate adaptation policy need probabilities ?, *4*(March), 37–41.

Dore, M. H. I. (2005). Climate change and changes in global precipitation patterns: What do we know? *Environment International*, *31*(8), 1167–1181. https://doi.org/10.1016/j.envint.2005.03.004

Frei, C., Schöll, R., Fukutome, S., Schmidli, J., & Vidale, P. L. (2006). Future change of precipitation extremes in Europe: Intercomparison of scenarios from regional climate models. *Journal of Geophysical Research Atmospheres*, *111*(6). https://doi.org/10.1029/2005JD005965

Grubbs, F. E., & Beck, G. (2019). Extension of Sample Sizes and Percentage Points for Significance Tests of Outlying Observations. *American Society for Quality*, *14*(4), 847–854.

IPCC. (2013). Climae Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovern- mental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535 pp.

CoverJiang, T., Chen, Y. D., Xu, C. yu, Chen, X., Chen, X., & Singh, V. P. (2007). Comparison of hydrological impacts of climate change simulated by six hydrological models in the Dongjiang Basin, South China, *336*(3–4), 316–333. https://doi.org/10.1016/j.jhydrol.2007.01.010

Hosking, J.R.M. & Wallis, J. R. (1997). *Regional frequency analysis: an approach based on L-moments*. Cambridge, New York. Cambridge University Press.

Katz, R. W., Parlange, M. B., & Naveau, P. (1989). Statistics of extremes in climatology. *Advances in Water Resources*, *43*(1), 1–30. https://doi.org/10.1111/j.1467-9574.1989.tb01244.x

Kendon, E. J., Blenkinsop, S., & Fowler, H. J. (2018). When will we detect changes in short-duration precipitation extremes? *Journal of Climate*, *31*(7), 2945–2964. https://doi.org/10.1175/JCLI-D-17-0435.1

Kharin, V. V., Zwiers, F. W., Zhang, X., & Wehner, M. (2013). Changes in temperature and precipitation extremes in the CMIP5 ensemble. *Climatic Change*, *119*(2), 345–357. https://doi.org/10.1007/s10584-013-0705-8

Kundzewicz, Z. W., & Stakhiv, E. Z. (2010). Are climate models "ready for prime time" in water resources management applications, or is more research needed? *Hydrological Sciences Journal*, *55*(7), 1085–1089. https://doi.org/10.1080/02626667.2010.513211

Lenderink, G., & Van Meijgaard, E. (2008). Increase in hourly precipitation extremes beyond expectations from temperature changes. *Nature Geoscience*, *1*(8), 511–514. https://doi.org/10.1038/ngeo262

Li, H., Sheffield, J., Wood, E. (2010). Bias correction of monthly precipitation and temperature fields from intergovernmental Panel on Climate Change AR4 models using equidistant quantile matching. *Journal of Geophysical Research Atmospheres, 115(10).*

Mallakpour, I., & Villarini, G. (2015). The changing nature of flooding across the central United States. *Nature Climate Change*, *5*(3), 250–254. https://doi.org/10.1038/nclimate2516

Milly, P. C. D., Bentacourt, J., Falkenmark, M., Robert, M., Hirsch, R. M., Kundzewicz, Z. W., … Stouffer, R. J. (2008). Stationarity is dead: Whither water management? . *Science*, *319*(February), 573–574.

Mondal, A., & Mujumdar, P. P. (2015). Modeling non-stationarity in intensity, duration and frequency of extreme rainfall over India. *Journal of Hydrology*, *521*, 217–231. https://doi.org/10.1016/j.jhydrol.2014.11.071

Noaa. (2013). Precipitation-Frequency Atlas of the United States. *Precipitation-Frequency Atlas of the United States*, *8*.

O'Gorman, P. A., & Schneider, T. (2009). The physical basis for increases in precipitation extremes in simulations of 21st-century climate change. *Proceedings of the National Academy of Sciences*, *106*(35), 14773–14777. https://doi.org/10.1073/pnas.0907610106

Oki, T., & Kanae, S. (2006). Global hydrological cycles and world water resources. *Science*, *313*(5790), 1068–1072. https://doi.org/10.1126/science.1128845

Pendergrass, A. G. (2018). What precipitation is extreme? *Science*, *360*(6393), 1072–1073. https://doi.org/10.1126/science.aat1871

Pendergrass, A. G., & Knutti, R. (2018). The Uneven Nature of Daily Precipitation and Its Change. *Geophysical Research Letters*, *45*(21), 11,980-11,988. https://doi.org/10.1029/2018GL080298

Pfahl, S., O'Gorman, P. A., & Fischer, E. M. (2017). Understanding the regional pattern of projected future changes in extreme precipitation. *Nature Climate Change*, *7*(6), 423–427. https://doi.org/10.1038/nclimate3287

Prein, A. F., Rasmussen, R. M., Ikeda, K., Liu, C., Clark, M. P., & Holland, G. J. (2017). The future intensification of hourly precipitation extremes. *Nature Climate Change*, *7*(1), 48–52. https://doi.org/10.1038/nclimate3168

Rosbjerg, D. A. N., & Madsen, H. K. (1992). Prediction in Partial Duration Series With Generalized. *Water Resources Research*, *28*(11), 3001–3010.

Salas, J. D., Obeysekera, J., & Vogel, R. M. (2018). Techniques for assessing water infrastructure for nonstationary extreme events: a review. *Hydrological Sciences Journal*, *63*(3), 325–352. https://doi.org/10.1080/02626667.2018.1426858

Sarhadi, A., & Soulis, E. D. (2017). Time-varying extreme rainfall intensity-duration-frequency curves in a changing climate. *Geophysical Research Letters*, *44*(5), 2454–2463. https://doi.org/10.1002/2016GL072201

Sharma, A., Wasko, C., & Lettenmaier, D. P. (2018). If Precipitation Extremes Are Increasing, Why Aren't Floods? *Water Resources Research*, *54*(11), 8545–8551. https://doi.org/10.1029/2018WR023749

Solari, S., & Losada, M. A. (2012). A unified statistical model for hydrological variables including the selection of threshold for the peak over threshold method, *48*(October 2011), 1–15. https://doi.org/10.1029/2011WR011475

Šraj, M., Viglione, A., Parajka, J., & Blöschl, G. (2016). The influence of non-stationarity in extreme hydrological events on flood frequency estimation. *Journal of Hydrology and Hydromechanics*, *64*(4), 426–437. https://doi.org/10.1515/johh-2016-0032

Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, *93*(4), 485–498. https://doi.org/10.1175/BAMS-D-11-00094.1

U.S. Water Resources Council. (1982). *Guidelines for Determining Flood Flow Frequency*. *Bulletin 17B: Reston, Virginia, Hydrology Subcommittee, Office of Water Data Coordination, U.S. Geological Survey*. https://doi.org/http://dx.doi.org/10.3133/tm4–BXX/

Wang, G., Wang, D., Trenberth, K. E., Erfanian, A., Yu, M., Bosilovich, M. G., & Parr, D. T. (2017). The peak structure and future changes of the relationships between extreme precipitation and temperature. Nature Climate Change, 7(4), 268–274. https://doi.org/10.1038/nclimate3239

Yue, S., Pilon, P., & Cavadias, G. (2002). Power of the Mann-Kendall and Spearman's rho tests for detecting monotonic trends in hydrological series. *Journal of Hydrology*, *259*(1–4), 254–271. https://doi.org/10.1016/S0022-1694(01)00594-7